

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
3 January 2002 (03.01.2002)(10) International Publication Number  
WO 02/00856 A2

PCT

(51) International Patent Classification: C12N 9/00

M.; Kerschbaumstr. 99, B-8810 Lichterwende (BE);  
GEYSENS, Steven, C., J.; Steenbeekstraat 8, B-9770  
Kroisbuiten (BE).

(21) International Application Number: PCT/IB01/01514

(22) International Filing Date: 29 June 2001 (29.06.2001)

(74) Agent: PUST, Detlev, A.; Hoffmann - Biele, Ambellast-  
trasse 4, 81925 München (DE).

(25) Filing Language: English

English

(26) Publication Language: English

English

(30) Priority Data: 60215,676

30 June 2000 (30.06.2000) US

Published:

(71) Applicant: FLANDERS INTERUNIVERSITY INSTITUTE FOR BIOTECHNOLOGY (VIB) [BE/BH]; Rijksweg 120, B-9052 Zwijnaarde (BE).

— without international search report and to be republished upon receipt of that report

(72) Inventors: CONTRERAS, Roland; Molensstraat 53, B-9820 Merelbeke (BE); CALLEWAERT, Nico, L., For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Protein Glycosylation Modification in *Pichia Pastoris***Field of the Invention**

The present invention relates to methods and vectors useful for genetically modifying the glycosylation process in methylotrophic yeast strains for the purpose of producing glycoproteins with reduced glycosylation. The present invention further relates to methylotrophic yeast strains generated using the present methods and vectors, as well as glycoproteins produced from such genetically modified strains.

**10 Background of the Invention**

The methylotrophic yeasts including *Pichia pastoris* have been widely used for production of recombinant proteins of commercial or medical importance.

However, production and medical applications of some therapeutic glycoproteins can be hampered by the differences in the protein-linked carbohydrate biosynthesis between these yeasts and the target organism such as a mammalian subject.

Protein N-glycosylation originates in the endoplasmic reticulum (ER),

where an N-linked oligosaccharide (Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub>) assembled on dolichol (a lipid carrier intermediate) is transferred to the appropriate Asn of a nascent protein. This is

an event common to all eukaryotic N-linked glycoproteins. The three glucose residues

and one specific α-1,2-linked mannose residue are removed by specific glucosidases

and an α-1,2-mannosidase in the ER, resulting in the core oligosaccharide structure,

Man<sub>5</sub>GlcNAc<sub>2</sub>. The protein with this core sugar structure is transported to the Golgi apparatus where the sugar moiety undergoes various modifications. There are

significant differences in the modifications of the sugar chain in the Golgi apparatus between yeast and higher eukaryotes.

In mammalian cells, the modification of the sugar chain proceeds via 3

different pathways depending on the protein moiety to which it is added. That is, (1)

the core sugar chain does not change; (2) the core sugar chain is changed by adding

the N-acetylglucosamine-1-phosphate moiety (GlcNAc-1-P) in UDP-N-acetyl

(54) Title: PROTEIN GLYCOSYLATION MODIFICATION IN *PICHTA PASTORIS*

(57) Abstract: The present invention provides genetically engineered strains of *Pichia* capable of producing proteins with reduced glycosylation. In particular, the genetically engineered strains of the present invention are capable of expressing either or both of an α-1,2-mannosidase and glucosidase II. The genetically engineered strains of the present invention can be further modified such that the OCH1 gene is disrupted. Methods of producing glycoproteins with reduced glycosylation using such genetically engineered strains of *Pichia* are also provided.

BEST AVAILABLE COPY

WO 02/00856 A2

glucosamine (UDP-GlcNAc) to the 6-position of mannose in the core sugar chain, followed by removing the GlcNAc moiety to form an acidic sugar chain in the glycoprotein; or (3) the core sugar chain is first converted into Man<sub>5</sub>GlcNAc<sub>2</sub> by removing 3 mannose residues with mannosidase I; Man<sub>5</sub>GlcNAc<sub>2</sub> is further modified by adding GlcNAc and removing 2 more mannose residues, followed by sequentially adding GlcNAc, galactose (Gal), and N-acetylneuraminic acid (also called sialic acid (NeuNAc)) to form various hybrid or complex sugar chains (R. Kornfeld and S. Kornfeld, *Ann. Rev. Biochem.* 54: 631-664, 1985; Cliba et al *J. Biol. Chem.* 273: 26298-26304, 1998).

In yeast, the modification of the sugar chain in the Golgi involves a series of additions of mannose residues by different mannosyltransferases ("outer chain" glycosylation). The structure of the outer chain glycosylation is specific to the organisms, typically with more than 50 mannose residues in *S. cerevisiae*, and most commonly with structures smaller than Man<sub>14</sub>GlcNAc<sub>2</sub> in *Pichia pastoris*. This yeast-specific outer chain glycosylation of the high mannose type is also denoted hyperglycosylation.

Hyperglycosylation is often undesired since it leads to heterogeneity of a recombinant protein product in both carbohydrate composition and molecular weight, which may complicate the protein purification. The specific activity (units/weight) of hyperglycosylated enzymes may be lowered by the increased portion of carbohydrate. In addition, the outer chain glycosylation is strongly immunogenic which is undesirable in a therapeutic application. Moreover, the large outer chain sugar can mask the immunogenic determinants of a therapeutic protein. For example, the influenza neuraminidase (NA) expressed in *P. pastoris* is glycosylated with N-glycans containing up to 30-40 mannose residues. The hyperglycosylated NA has a reduced immunogenicity in mice, as the variable and immunodominant surface loops on top of the NA molecule are masked by the N-glycans (Martinet et al. *Eur J. Biochem.* 247: 332-338, 1997).

Therefore, it is desirable to genetically engineer methylotrophic yeast strains in which glycosylation of proteins can be manipulated and from which recombinant proteins can be produced that would not be compromised in structure or function by large N-glycan side chains.

#### Summary of the Invention

The present invention is directed to methods and vectors useful for genetically modifying the glycosylation process in methylotrophic yeast strains to produce glycoproteins with reduced glycosylation. Methylotrophic yeast strains generated using the present methods and vectors, as well as glycoproteins produced from such genetically modified strains are also provided.

In one embodiment, the present invention provides vectors useful for making genetically engineered methylotrophic yeast strains which are capable of producing glycoproteins with reduced glycosylation.

In one aspect, the present invention provides "knock-in" vectors which are capable of expressing in a methylotrophic yeast strain one or more proteins whose enzymatic activities lead to a reduction of glycosylation in glycoproteins produced by the methylotrophic yeast strain.

In a preferred embodiment, the knock-in vectors of the present invention include a nucleotide sequence coding for an  $\alpha$ -1,2-mannosidase or a functional part thereof and are capable of expressing the  $\alpha$ -1,2-mannosidase or the functional part in a methylotrophic yeast strain. A preferred nucleotide sequence is a nucleotide sequence encoding the  $\alpha$ -1,2-mannosidase of a fungal species, and more preferably, *Trichoderma reesei*. Preferably, the  $\alpha$ -1,2-mannosidase expression vector is engineered such that the  $\alpha$ -1,2-mannosidase or a functional part thereof expressed from the vector includes an ER-retention signal. A preferred ER-retention signal is HDEL. The  $\alpha$ -1,2-mannosidase coding sequence can be operable linked to a constitutive or inducible promoter, and a 3' termination sequence. The vectors can be

integrative vectors or replicative vectors. Particularly preferred  $\alpha$ -1,2-mannosidase expression vectors include pGAPZMFMManHDEL, pGAPZMFMManMycHDEL, pPICZBMFMManMycHDEL, pGAPZmManHDEL, pGAPZmMycManHDEL, pPIC9mMycManHDEL and pGAPZmMycManHDEL.

5 In another preferred embodiment, the knock-in vectors of the present invention include a sequence coding for a glucosidase II or a functional part thereof and are capable of expressing the glucosidase II or the functional part in a methylotrophic yeast strain. A preferred nucleotide sequence is a nucleotide sequence encoding the glucosidase II of a fungal species, and more preferably, *Saccharomyces cerevisiae*. Preferably, the glucosidase II expression vector is engineered such that the glucosidase II or a functional part thereof expressed from the vector includes an ER-retention signal. A preferred ER-retention signal is HDEL. The glucosidase II coding sequence can be operable linked to a constitutive or inducible promoter, and a 3' termination sequence. The vectors can be integrative vectors or replicative vectors.

15 Particularly preferred glucosidase II expression vectors include pGAPZAGLSII, pPICZAGLSII, pAOX2ZAGLSII, pYPTIZAGLSII, pGAPADEglsII, pICADEglsII, pAOX2ADEglsII, pYPTIADEglsII, pGAPZAGlsIHDEL and pGAPADEglsIHDEL.

Expression vectors which include both of an  $\alpha$ -1,2-mannosidase expression unit and a glucosidase II expression unit are also provided by the present invention.

20 In another aspect, the present invention provides "knock-out" vectors which, when introduced into a methylotrophic yeast strain, inactivate or disrupt a gene thereby facilitating the reduction in the glycosylation of glycoproteins produced in the methylotrophic yeast strain.

25 In one embodiment, the present invention provides a "knock-out" vector which, when introduced into a methylotrophic yeast strain, inactivates or disrupts the Och1 gene. A preferred Och1 knock-out vector is pBLURA5'pOCHI.

Still another embodiment of the present invention provides vectors which include both a knock-in unit and a knock-out unit.

Furthermore, any of the knock-in or knock-out vectors of the present invention can also include a nucleotide sequence capable of expressing a heterologous protein of interest in a methylotrophic yeast.

5 Another embodiment of the present invention provides methods of modifying the glycosylation in a methylotrophic yeast by transforming the yeast with one or more vectors of the present invention.

10 Strains of a methylotrophic yeast which can be modified using the present methods include, but are not limited to, yeast strains capable of growth on methanol such as yeasts of the genera *Candida*, *Hansenula*, *Torulopsis*, and *Pichia*. Preferred methylotrophic yeasts are of the genus *Pichia*. Especially preferred are *Pichia pastoris* strains GSI15 (NRRL Y-15851), GSI90 (NRRL Y-18014), PPF1 (NRRL Y-18017), PPY120H, yGC4, and strains derived therefrom. Methylotrophic yeast strains which can be modified using the present methods also include those methylotrophic yeast strains which have been engineered to express one or more heterologous proteins of interest. The glycosylation on the heterologous proteins expressed from these previously genetically engineered strains can be reduced by transforming such strains with one or more of the vectors of the present invention.

15 Methylotrophic yeast strains which are modified by practicing the present methods are provided in another embodiment of the present invention.

20 A further aspect of the present invention is directed to methods of producing glycoproteins with a reduced glycosylation.

25 In accordance with such methods, a nucleotide sequence capable of expressing a glycoprotein can be introduced into a methylotrophic yeast strain which has previously been transformed with one or more of the vectors of the present invention. Alternatively, a methylotrophic yeast strain which has been genetically engineered to express a glycoprotein can be transformed with one or more of the vectors of the present invention. Moreover, if a methylotrophic yeast strain is not transformed with a nucleotide sequence encoding a glycoprotein of interest or any of the vectors of the present invention, such yeast strain can be transformed, either

consecutively or simultaneously, with both a nucleotide sequence capable of expressing the glycoprotein and one or more vectors of the present invention.

Additionally, a methylotrophic yeast strain can be transformed with one or more of the present knock-in and/or knock-out vectors which also include a nucleotide sequence capable of expressing a glycoprotein in the methylotrophic yeast strain.

Glycoproteins products produced by using the methods of the present invention, i.e., glycoproteins with reduced N-glycosylation, are also part of the present invention.

Kits which include one or more of the vectors of the present invention, or one or more strains modified to produce glycoproteins with reduced glycosylation, are also provided.

#### Brief Description of the Drawings

Figure 1 depicts vectors carrying an HDEL-tagged *Trichoderma reesei*  $\alpha$ -

1,2-mannosidase expression cassette and describes the way in which these vectors were constructed according to methods known in the art. Abbreviations used

throughout construction schemes: 5' AOX1 or AOX1 P: *Pichia pastoris* AOX1 promoter sequence; Amp R: ampicillin resistance gene; ColEI1: ColEI1 origin of replication; 3' AOX1: 3' sequences of the *Pichia pastoris* AOX1 gene; HIS4: HIS4

gene of *Pichia pastoris*. AOX TT: transcription terminator sequence of the *Pichia pastoris* AOX1 gene; ORF: open reading frame; S: secretion signal; P TEF1: the promoter sequence of the *Saccharomyces cerevisiae* transcription elongation factor 1 gene; P EM7: synthetic constitutive prokaryotic promoter EM7; Zeocin: Zeocin resistance gene; CYC1 TT: 3' end of the *S. cerevisiae* CYC1 gene; GAP: promoter

sequence of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene;

PpURA3: *Pichia pastoris* URA3 gene. As can be seen in this figure, the *Trichoderma reesei*  $\alpha$ -1,2-mannosidase was operably linked to the coding sequence for the *S.*

*cerevisiae*  $\alpha$ -mating factor secretion signal sequence and further operably linked at the 3' terminus of the coding sequence to the coding sequence for an HDEL peptide. The

whole fusion construct was operably linked to either the *P. pastoris* AOX1 promoter (in pPIC9MFMHDEL) or to the *P. pastoris* GAP promoter (in pGAPZMFMHDEL).

Figure 2 depicts vectors carrying an HDEL-tagged *Mus musculus*  $\alpha$ -1,2-

mannosidase IB expression cassette and describes the way in which these vectors were constructed according to methods known in the art. As can be seen in this figure, the

catalytic domain of the *Mus musculus*  $\alpha$ -1,2-mannosidase IB was operably linked to

the coding sequence for the *S. cerevisiae*  $\alpha$ -mating factor secretion signal sequence

and further operably linked at the 3' terminus of the coding sequence to the coding

sequence for an HDEL peptide. The whole fusion construct was operably linked to

either the *P. pastoris* AOX1 promoter (in pPIC9mManHDEL) or to the *P. pastoris*

GAP promoter (in pGAPZmManHDEL). Furthermore, variants of the expression

cassette were made in which the coding sequence for a cMyc epitope tag was inserted

between the coding sequence for the *S. cerevisiae*  $\alpha$ -Mating Factor secretion signal

sequence and the coding sequence for the catalytic domain of the *Mus musculus*  $\alpha$ -1,2-

mannosidase IB. This expression cassette was also operably linked to either the *P.*

*pastoris* AOX1 promoter (in pPIC9mMycManHDEL) or to the *P. pastoris* GAP

promotor (in pGAPZmMycManHDEL).

Figure 3 depicts vectors carrying a MycHDEL tagged *Trichoderma reesei*

$\alpha$ -1,2-mannosidase and the way in which these vectors were obtained. The resulting

fusion construction was again operably linked to either the *P. pastoris* AOX1

promoter (in pPICZBMFMHDEL) or to the *P. pastoris* GAP promoter (in

pGAPZMFMHDEL).

Figure 4 demonstrates the intracellular localization of the MycHDEL-

tagged *Trichoderma reesei*  $\alpha$ -1,2-mannosidase and indicates ER-targeting by

immunofluorescence analysis. Panel A Western blotting. Yeast strains were grown in

10 ml YPG cultures to an OD<sub>600</sub>=10, diluted fivefold and grown in YPM for 48 h.

1/50th of the culture medium and 1/65th of the cells were analysed by SDS-PAGE and

Western blotting with the mouse monoclonal 9E10 anti-Myc antibody. The position of molecular weight marker proteins are indicated with arrows. Lanes 1-5: cellular lysates. 1,2: pGAPZMFManMycHDEL transformants. 3: non-transformed PPY12OH (negative control). 4,5: pPICZBMFManMycHDEL transformants. Lanes 6-10: culture media. 6: non transformed PPY12OH (negative control). 7,8: pGAPZMFManMycHDEL transformants. 9,10: pPICZBMFManMycHDEL transformants. *Panel B* Immunofluorescence microscopy. 1: phase contrast image of a *P. pastoris* cell (strain PPY12OH transformed with pGAPZMFManHDEL) at 1000x magnification. The nucleus is visible as an ellipse in the lower right quadrant of the cell. 2: same cell as in 1, but in fluorescence microscopy mode to show localization of the *T. reesei* mannosidase-Myc-HDEL protein. The protein is mainly localized in a circular distribution around the nucleus (nuclear envelope), which is typical for an endoplasmic reticulum steady-state distribution. 3: phase contrast image of a *P. pastoris* cell (strain PPY12OH transformed with pGAPZMFManHDEL) at 1000x magnification. 4: same cell in fluorescence microscopy to show localization of the Golgi marker protein OCH1-HA in *P. pastoris* strain PPY12OH. The dot-like distribution throughout the cytoplasm, with 3-4 dots per cell is typical for cis-Golgi distribution in *P. pastoris*.

5

Figure 5 depicts the co-sedimentation of mannosidase-MycHDEL with Protein Disulfide Isomerase in sucrose density gradient centrifugation. The top panel shows the distribution over the different fractions of the sucrose gradient of the OCH1-HA Golgi marker protein. The middle panel shows this distribution for the Protein Disulfide Isomerase endoplasmic reticulum marker protein. Finally, the bottom panel shows the distribution of the MycHDEL-tagged *Trichoderma reesei*  $\alpha$ -1,2-mannosidase over the same fractions. It is concluded that the mannosidase-MycHDEL almost exactly matches the distribution of the ER marker PDI and thus mainly resides in the ER of the *Pichia pastoris* yeast cells.

15

Figure 6 depicts the N-glycan analysis of *Trypanosoma cruzi trans-*

sialidase coexpressed with *Trichoderma reesei* mannosidase-HDEL. Panel A: malto-

25

oligosaccharide size reference ladder. Sizes of the glycans are expressed in Glucose Units (GU) by comparison of their electrophoretic mobility to the mobility of these malto-oligosaccharides. Panel B: N-glycans derived from recombinant *Trypanosoma cruzi trans*-sialidase expressed in *Pichia pastoris*. The peak at GU=9,2 corresponds to Man<sub>5</sub>GlcNAc<sub>2</sub>. Panel C: same analytes as panel 2, but after overnight treatment with 3U/ml purified recombinant *T. reesei*  $\alpha$ -1,2-mannosidase. Panel D: N-glycans derived from recombinant *trans*-sialidase co-expressed in *Pichia pastoris* with *T. reesei* mannosidase-HDEL (under control of the GAP promoter). The peak at GU=7,6 corresponds to the Man<sub>5</sub>GlcNAc<sub>2</sub> peak in the profile of RNase B (Panel F). Panel E: same analytes as panel D, but after overnight treatment with 3 mU/ml purified recombinant *T. reesei*  $\alpha$ -1,2-mannosidase. Panel F: N-glycans derived from bovine RNase B. These glycans consist of Man<sub>5</sub>GlcNAc<sub>2</sub> to Man<sub>9</sub>GlcNAc<sub>2</sub>. Different isomers are resolved, accounting for the number of peaks for Man<sub>7</sub>GlcNAc<sub>2</sub>.

10

Figure 7 depicts the processing of influenza haemagglutinin N-glycans by HDEL-tagged *Trichoderma reesei*  $\alpha$ -1,2-mannosidase and the HDEL-tagged catalytic domain of murine  $\alpha$ -1,2-mannosidase IB. The Man<sub>5</sub>GlcNAc<sub>2</sub> reference oligosaccharide runs at scan 1850 in this analysis (not shown). Panel 1: malto-

15

oligosaccharide size reference ladder. Panel 2: N-glycans derived from recombinant influenza haemagglutinin expressed in *Pichia pastoris*. The peak at scan 2250 corresponds to Man<sub>5</sub>GlcNAc<sub>2</sub>. Panel 3: N-glycans derived from recombinant haemagglutinin co-expressed in *Pichia pastoris* with *T. reesei* mannosidase-HDEL (under control of the GAP promoter). The peak at scan 1950 corresponds to Man<sub>5</sub>GlcNAc<sub>2</sub>. Panel 4: Same analytes as for panel 3, but after overnight treatment with 30 mU purified recombinant *T. reesei*  $\alpha$ -1,2-mannosidase. Panel 5: N-glycans derived from recombinant haemagglutinin co-expressed in *Pichia pastoris* with mouse mannosidase IB-HDEL (under control of the GAP promoter). Panel 6: same analytes as for panel 5, but after overnight treatment with 30 mU purified recombinant *T. reesei*  $\alpha$ -1,2-mannosidase.

20

25

Figure 8 graphically depicts vector pBLURAS<sup>+</sup>PpOCH1 and the way in which it was constructed.

Figure 9 depicts the scheme for disrupting the *Pichia pastoris* OCH1 gene by single homologous recombination using pBLURAS<sup>+</sup>PpOCH1.

Figure 10 depicts the cell wall glycoprotein N-glycan analysis of the OCH1-inactivated clone and three clones derived from this OCH1-inactivated clone by transformation with pGAPZMFManHDEL. Panel 1 shows the analysis of a mixture of malto-oligosaccharides, the degree of polymerisation of which is given by the numbers on the very top of the figure. This analysis serves as a size reference for the other panels. On the vertical axis of all panels, peak intensity in relative fluorescence units is given. Panel 2-6: analysis of the cell wall glycoprotein N-glycans of the following strains: Panel 2, non-engineered *P. pastoris* strain yGCA; Panel 3, yGCA transformed with pBLURAS<sup>+</sup>PpOch1; 4-6, three clones of the strain of Panel 3, supplementarily transformed with pGAPZMFManHDEL. Panel 7: the N-glycans derived from bovine RNaseB, consisting of a mixture of Man<sub>5</sub>-GlcNAc<sub>2</sub>. As can be seen from comparison between panel 2 and 3 and reference to panel 7, transformation with pBLURAS<sup>+</sup>PpOch1 leads to a strongly increased abundance of the Man<sub>5</sub>GlcNAc<sub>2</sub> substrate N-glycan (named peak 1 in Panel 2) of OCH1p. Peak 2 represents the Man<sub>5</sub>GlcNAc<sub>2</sub> product of OCH1p. Furthermore, upon supplementary transformation of pGAPZMFManHDEL, the major glycan on the cell wall glycoproteins of three independent clones is the Man<sub>5</sub>GlcNAc<sub>2</sub> end product (peak 3 in panel 4) of *T. reesei*  $\alpha$ -1,2-mannosidase digestion of the Man<sub>5</sub>GlcNAc<sub>2</sub> substrate.

Figure 11 depicts the analysis of exactly the same glycan mixtures as in Figure 10, but after an *in vitro* digest with 3mU/ml purified *Trichoderma reesei*  $\alpha$ -1,2-mannosidase, overnight in 20 mM sodium acetate pH=5.0. Axis assignment is the same as in Figure 10. More Man<sub>5</sub>GlcNAc<sub>2</sub> is formed in the pBLURAS<sup>+</sup>PpOch1 transformed strain (Panel 3) than in the parent strain (Panel 2). Peaks in all panels before scan 3900 come from contaminants and should be ignored in the analysis.

Figure 12 depicts the expression vector pGAPZAGLSII (SEQ ID NO: 18). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: bacterial origin of replication. GAP: promotor of the *P. pastoris* GAP gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene

Figure 13 depicts the expression vector pAOXZAGLSII (SEQ ID NO: 16). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: bacterial origin of replication. AOX2 P: promotor of the *P. pastoris* AOX2 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene

Figure 14 depicts the expression vector pPICZAGLSII (SEQ ID NO: 20).

P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: origin of replication. AOX1 P: promotor of the *P. pastoris* AOX1 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene

Figure 15 depicts the expression vector pYPT1ZAGLSII (SEQ ID NO: 22). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: origin of replication. P YPT1: promotor of the *P. pastoris* YPT1 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 16 depicts the expression vector pGAPADE1glsII (SEQ ID NO: 19). Amp R: Ampicillin resistance marker gene. ADE1: *P. pastoris* ADE1 selection

marker gene. GAP: promotor of the *P. pastoris* GAP gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene

Figure 17 depicts the expression vector pAOX2ADE1glsII (SEQ ID NO:

17). Amp R: Ampicillin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. AOX2 P: promotor of the *P. pastoris* AOX2 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 18 depicts the expression vector pPICADE1glsII (SEQ ID NO: 21).

Amp R: Ampicillin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. AOX1 P: promotor of the *P. pastoris* AOX1 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 19 depicts the expression vector pYPT1ADE1glsII (SEQ ID NO:

23). Amp R: Ampicillin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. P YPT1: promotor of the *P. pastoris* YPT1 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 20 depicts the expression vector pGAPZAglsIIHDEL (SEQ ID NO:

24). Amp R: Ampicillin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. GAP: promotor of the *P. pastoris* GAP gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 21 depicts the expression vector pGAPADE1glsIIHDEL (SEQ ID

NO: 25). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: bacterial origin of replication. GAP: promotor of the *P. pastoris* GAP gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 22 depicts the test of the GLSII activity assay using a commercially available yeast alpha-glucosidase (Sigma: Cat. No. G-5003). The assay mixture contains phosphate-citrate buffer pH 6.8, mannose, 2-deoxy-D-glucose, the substrate

4-methylumbelliferyl-alpha-D-glucopyranoside and alpha-glucosidase from Sigma. 1: assay mixture illuminated with UV-light after overnight incubation at 37 ° C; 2: same as 1, but this time, the assay mixture lacks the alpha-glucosidase; 3: same as 1, but this time, the assay mixture lacks the substrate.

Figure 23 depicts the results of the activity of recombinantly expressed GLSII from *Pichia pastoris*. All assay mixtures were incubated overnight at 37 ° C and afterwards illuminated with UV-light. 1: assay with yeast alpha-glucosidase (Sigma: Cat. No. G-5003); 2: assay with the purified medium of strain 18 (PPY12-OH transformed with pGAPZAglsII); 3: assay with purified medium of the WT PPY12-OH strain; 4: assay with the purified medium of strain H3 (PPY12-OH transformed with pGAPZAglsIIHDEL).

#### Detailed Description of the Invention

It has been established that the majority of N-glycans on glycoproteins leaving the endoplasmic reticulum (ER) of *Pichia* have the core Man<sub>9</sub>GlcNAc<sub>2</sub> oligosaccharide structure. After the proteins are transported from the ER to the Golgi apparatus, additional mannose residues are added to this core sugar moiety by different mannosyltransferases, resulting in glycoproteins with large mannose side chains. Such hyperglycosylation of recombinant glycoproteins is undesirable in many instances. Accordingly, the present invention provides methods and vectors for genetically modifying methylotrophic yeast strains to produce glycoproteins with reduced glycosylation. Methylotrophic yeast strains generated using the present methods and vectors, as well as glycoproteins produced from such genetically modified strains are also provided.

In one embodiment, the present invention provides vectors useful for genetically modifying methylotrophic yeast strains to produce glycoproteins with reduced glycosylation.

In one aspect, the present invention provides "knock-in" vectors which are capable of expressing in a methylotrophic yeast strain one or more proteins whose enzymatic activities lead to a reduction of glycosylation in glycoproteins produced by the methylotrophic yeast strain. According to the present invention, such proteins include, e.g., an  $\alpha$ -1,2-mannosidase, a glucosidase II, or functional parts thereof.

In a preferred embodiment, the vectors of the present invention include a sequence coding for an  $\alpha$ -1,2-mannosidase or a functional part thereof and are capable of expressing the  $\alpha$ -1,2-mannosidase or the functional part in a methylotrophic yeast strain.

An  $\alpha$ -1,2-mannosidase cleaves the  $\alpha$ -1,2-linked mannose residues at the non-reducing ends of  $\text{Man}_5\text{GlcNAc}_2$ , and converts this core oligosaccharide on glycoproteins to  $\text{Man}_3\text{GlcNAc}_2$ . *In vitro*,  $\text{Man}_5\text{GlcNAc}_2$  is a very poor substrate for any *Pichia* Golgi mannosyltransferase, i.e., mannose residues can not be added to this sugar structure. On the other hand,  $\text{Man}_5\text{GlcNAc}_2$  is the acceptor substrate for the mammalian N-acetylglucosaminyl-transferase I and is an intermediate for the hybrid- and complex-type sugar chains characteristic of mammalian glycoproteins. Thus, by way of introducing an  $\alpha$ -1,2-mannosidase into methylotrophic yeasts such as *Pichia*, glycoproteins with reduced mannose content can be produced.

According to the present invention, the nucleotide sequence encoding an  $\alpha$ -1,2-mannosidase for use in the expression vector of the present invention can derive from any species. A number of  $\alpha$ -1,2-mannosidase genes have been cloned and are available to those skilled in the art, including mammalian genes encoding, e.g., a murine  $\alpha$ -1,2-mannosidase (Herscovics et al. *J. Biol. Chem.* 269: 9864-9871, 1994), a rabbit  $\alpha$ -1,2-mannosidase (Lal et al. *J. Biol. Chem.* 269: 9872-9881, 1994) or a human  $\alpha$ -1,2-mannosidase (Tremblay et al. *Glycobiology* 8: 585-595, 1998), as well as fungal

genes encoding, e.g., an *Aspergillus*  $\alpha$ -1,2-mannosidase (*msdS* gene), a *Trichoderma reesei*  $\alpha$ -1,2-mannosidase (Maras et al. *J. Biotechnol.* 77: 255-263, 2000), or a *Saccharomyces cerevisiae*  $\alpha$ -1,2-mannosidase. Protein sequence analysis has revealed a high degree of conservation among the eukaryotic  $\alpha$ -1,2-mannosidases identified so far.

Preferably, the nucleotide sequence for use in the present vectors encodes a fungal  $\alpha$ -1,2-mannosidase, more preferably, a *Trichoderma reesei*  $\alpha$ -1,2-mannosidase, and more particularly, the *Trichoderma reesei*  $\alpha$ -1,2-mannosidase described by Maras et al. *J. Biotechnol.* 77: 255-63 (2000).

According to the present invention, the nucleotide sequence can also code for only a functional part of an  $\alpha$ -1,2-mannosidase.

By "functional part" is meant a polypeptide fragment of an  $\alpha$ -1,2-mannosidase which substantially retains the enzymatic activity of the full-length protein. By "substantially" is meant at least about 40%, or preferably, at least 50% or more of the enzymatic activity of the full-length  $\alpha$ -1,2-mannosidase is retained. For example, as illustrated by the present invention, the catalytic domain of the murine  $\alpha$ -1,2-mannosidase IB constitutes a "functional part" of the murine  $\alpha$ -1,2-mannosidase IB. Those skilled in the art can readily identify and make functional parts of an  $\alpha$ -1,2-mannosidase using a combination of techniques known in the art. Predictions of the portions of an  $\alpha$ -1,2-mannosidase essential to or sufficient to confer the enzymatic activity can be made based on analysis of the protein sequence. The activity of a portion of an  $\alpha$ -1,2-mannosidase of interest, expressed and purified from an appropriate expression system, can be verified using *in vitro* or *in vivo* assays described hereinbelow.

In accordance with the present invention, an  $\alpha$ -1,2-mannosidase or a functional part thereof expressed in a methylotrophic yeast strain preferably is targeted to a site in the secretory pathway where  $\text{Man}_5\text{GlcNAc}_2$  (the substrate of  $\alpha$ -1,2-

mannosidase) is already formed on a glycoprotein, but has not reached a Golgi glycosyltransferase which elongates the sugar chain with additional mannose residues.

Accordingly, in a preferred embodiment of the present invention, the  $\alpha$ -1,2-mannosidase expression vector is engineered as such that the  $\alpha$ -1,2-mannosidase or a functional part thereof expressed from the vector includes an ER-retention signal.

"An ER retention signal" refers to a peptide sequence which directs a protein having such peptide sequence to be transported to and retained in the ER. Such ER retention sequences are often found in proteins that reside and function in the ER.

Multiple choices of ER retention signals are available to those skilled in the art, e.g., the first 21 amino acid residues of the *S. cerevisiae* ER protein MNS1 (Martinet et al. *Biotechnology Letters* 20: 1171-1177, 1998). A preferred ER retention signal for use in the present invention is peptide HDEL (SEQ ID NO: 1). The HDEL peptide sequence, found in the C-terminus of a number of yeast proteins, acts as a retention/retrieval signal for the ER (Pelham *EMBO J.* 7: 913-918, 1988). Proteins with an HDEL sequence are bound by a membrane-bound receptor (Erd2p) and then enter a retrograde transport pathway for return to the ER from the Golgi apparatus. According to the present invention, an ER retention signal can be placed anywhere in the protein sequence of an  $\alpha$ -1,2-mannosidase, but preferably at the C-terminus of the  $\alpha$ -1,2-mannosidase.

The  $\alpha$ -1,2-mannosidase for use in the present invention can be further modified, e.g., by insertion of an epitope tag to which antibodies are available, such as Myc, HA, FLAG and His6 tags well-known in the art. An epitope-tagged  $\alpha$ -1,2-mannosidase can be conveniently purified, or monitored for both expression and intracellular localization.

An ER retention signal and an epitope tag can be readily introduced into a protein of interest by inserting nucleotide sequences coding for such signal or tag into

the nucleotide sequence encoding the protein of interest, using any of the molecular biology techniques known in the art.

In another preferred embodiment, the vectors of the present invention include a sequence coding for a glucosidase II or a functional part thereof and are capable of expressing the glucosidase II or the functional part in the methylotrophic yeast strain.

It has been established that the initial N-linked oligosaccharide (Glc<sub>2</sub>Man<sub>9</sub>GlcNAc<sub>2</sub>), transferred in the ER onto a protein, is cleaved in the ER by specific glucosidases to remove the glucose residues, and by a mannosidase to remove one specific  $\alpha$ -1,2-linked mannose. It has been observed by the present inventors that some recombinant proteins expressed in *Pichia* have residual glucose residues on the sugar moiety when such proteins leave the ER for the Golgi apparatus. The residual glucose molecules present on the sugar structure prevent the complete digestion of the sugar moiety by an  $\alpha$ -1,2-mannosidase, and the introduction of an exogenous glucosidase can facilitate the removal of these glucose residues.

According to the present invention, the nucleotide sequence encoding a glucosidase II can derive from any species. Glucosidase II genes have been cloned from a number of mammalian species including rat, mouse, pig and human. The glucosidase II protein from these mammalian species consists of an alpha and a beta subunit. The alpha subunit is about 110 kDa and contains the catalytic activity of the enzyme, while the beta subunit has a C-terminal HDEL ER-retention sequence and is believed to be important for the ER localization of the enzyme. The glucosidase II gene from *S. cerevisiae* has also been cloned (ORF YBR229c, located on chromosome II). This gene encodes a protein of about 110 kDa, which shows a high degree of homology to the mammalian alpha subunits.

A preferred glucosidase II gene for use in the present vectors is from a fungal species such as *Pichia pastoris* and *S. cerevisiae*. An example of a fungal glucosidase II gene is the *S. cerevisiae* glucosidase II alpha subunit gene.

According to the present invention, the nucleotide sequence can also encode only a functional part of a glucosidase II. By "functional part" is meant a polypeptide fragment of a glucosidase II which substantially retains the enzymatic activity of the full-length protein. By "substantially" is meant at least about 40%, or preferably, at least 50% or more of the enzymatic activity of the full-length glucosidase II is retained. Functional parts of a glucosidase II can be identified and made by those skilled in the art using a variety of techniques known in the art.

In a preferred embodiment of the present invention, the glucosidase II protein is engineered to include an ER retention signal such that the protein expressed in a methylotrophic yeast strain is targeted to the ER and retains therein for function. ER retention signals are as described hereinabove, e.g., the HDEL peptide sequence.

The glucosidase II for use in the present invention can be further modified, e.g., by insertion of an epitope tag to which antibodies are available, such as Myc, HA, FLAG, and His6 tag, which are well-known in the art.

According to the present invention, the "knock-in" vectors can include either or both of an  $\alpha$ -1,2-mannosidase coding sequence and a glucosidase II coding sequence.

Further according to the present invention, the nucleotide sequence coding for the enzyme to be expressed (e.g., an  $\alpha$ -1,2-mannosidase or a functional part thereof, or a glucosidase II or a functional part thereof) can be placed in an operable linkage to a promoter and a 3' termination sequence.

Promoters appropriate for expression of a protein in a methylotrophic yeast can include both constitutive promoters and inducible promoters. Constitutive promoters include e.g., the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase promoter ("the GAP promoter"). Examples of inducible promoters include, e.g., the *Pichia pastoris* alcohol oxidase I promoter ("the AOXI promoter") (U.S. Patent No. 4,855,231), or the *Pichia pastoris* formaldehyde dehydrogenase promoter ("the FLD promoter") (Shen et al. Gene 216: 93-102, 1998).

3' termination sequences are sequences 3' to the stop codon of a structural gene which function to stabilize the mRNA transcription product of the gene to which the sequence is operably linked, such as sequences which elicit polyadenylation. 3' termination sequences can be obtained from *Pichia* or other methylotrophic yeast.

Examples of *Pichia pastoris* 3' termination sequences useful for the practice of the present invention include termination sequences from the *AOXI* gene, *p40* gene, *HIS4* gene and *FLD1* gene.

The vectors of the present invention preferably contain a selectable marker gene. The selectable marker may be any gene which confers a selectable phenotype upon a methylotrophic yeast strain and allows transformed cells to be identified and selected from untransformed cells. The selectable marker system may include an auxotrophic mutant methylotrophic yeast strain and a wild type gene which complements the host's defect. Examples of such systems include the *Saccharomyces cerevisiae* or *Pichia pastoris HIS4* gene which may be used to complement *his4* *Pichia* strains, or the *S. cerevisiae* or *Pichia pastoris ARG4* gene which may be used to complement *Pichia pastoris arg* mutants. Other selectable marker genes which function in *Pichia pastoris* include the *Zeo<sup>R</sup>* gene, the *G418<sup>R</sup>* gene, and the like.

The vectors of the present invention can also include an autonomous replication sequence (ARS). For example, U.S. Patent No. 4,837,148 describes autonomous replication sequences which provide a suitable means for maintaining plasmids in *Pichia pastoris*. The disclosure of U.S. Patent No. 4,837,148 is incorporated herein by reference.

The vectors can also contain selectable marker genes which function in bacteria, as well as sequences responsible for replication and extrachromosomal maintenance in bacteria. Examples of bacterial selectable marker genes include ampicillin resistance (*Amp<sup>r</sup>*), tetracycline resistance (*Tet<sup>r</sup>*), neomycin resistance, hygromycin resistance, and zeocin resistance (*Zeo<sup>R</sup>*) genes.

According to the present invention, the nucleotide sequence encoding the protein to be expressed in a methylotrophic yeast can be placed in an integrative vector or a replicative vector (such as a replicating circular plasmid).

Integrative vectors are disclosed, e.g., in U.S. Patent No. 4,882,279 which is incorporated herein by reference. Integrative vectors generally include a serially arranged sequence of at least a first insertable DNA fragment, a selectable marker gene, and a second insertable DNA fragment. The first and second insertable DNA fragments are each about 200 nucleotides in length and have nucleotide sequences which are homologous to portions of the genomic DNA of the species to be transformed. A nucleotide sequence containing a structural gene of interest for expression is inserted in this vector between the first and second insertable DNA fragments whether before or after the marker gene. Integrative vectors can be linearized prior to yeast transformation to facilitate the integration of the nucleotide sequence of interest into the host cell genome.

Replicative and integrative vectors carrying either or both of an  $\alpha$ -1,2-mannosidase coding sequence or a glucosidase II coding sequence can be constructed by standard techniques known to one of ordinary skill in the art and found, for example, in Sambrook et al. (1989) in *Molecular Cloning: A Laboratory Manual*, or any of a myriad of laboratory manuals on recombinant DNA technology that are widely available.

Preferred vectors of the present invention carrying an  $\alpha$ -1,2-mannosidase expression sequence include pGAPZMFManHDEL, pGAPZMFManMycHDEL, pPICZBMFManMycHDEL, pGAPZmManHDEL, pGAPZmMycManHDEL, pPIC9mMycManHDEL and pGAPZmMycManHDEL, which are further described in the Examples hereinbelow.

Preferred vectors of the present invention carrying a glucosidase II expression sequence include pGAPZAGLSII, pPICZAGLSII, pAOX2ZAGLSII, pYPTIZAGLSII, pGAPADEIglslI, pPICADEIglslI, pAOXZADEIglslI,

pYPTIADEIglslI, pGAPZAglsIHDEL and pGAPADEIglslIHDEL, which are further described in the Examples hereinbelow.

In another aspect, the present invention provides "knock-out" vectors which, when introduced into a methylotrophic yeast strain, inactivate or disrupt a gene thereby facilitating the reduction in the glycosylation of glycoproteins produced in the methylotrophic yeast strain.

In one embodiment, the present invention provides a "knock-out" vector which, when introduced into a methylotrophic yeast strain, inactivates or disrupts the Och1 gene.

The *S. cerevisiae* OCH1 gene has been cloned (Nakayama et al. *EMBO J.* 11: 2511-2519, 1992). It encodes a membrane bound  $\alpha$ -1,6-mannosyltransferase, localized in the early Golgi complex, that is functional in the initiation of  $\alpha$ -1,6-polymannose outer chain addition to the N-linked core oligosaccharide (Man<sub>5</sub>GlcNAc<sub>2</sub> and Man<sub>9</sub>GlcNAc<sub>2</sub>) (Nakanishi-Shindo et al. *J. Biol. Chem.* 268: 26338-26345, 1993). A *Pichia* sequence has been described in Japanese Patent Application No. 07145005 that encodes a protein highly homologous to the *S. cerevisiae* OCH1. For purpose of the present invention, this sequence is denoted herein as "the *Pichia* OCH1 gene". Those skilled in the art can isolate the OCH1 genes from other methylotrophic yeasts using techniques well known in the art.

According to the present invention, a disruption in the OCH1 gene of a methylotrophic yeast can result in either the production of an inactive protein product or no product. The disruption may take the form of an insertion of a heterologous DNA sequence into the coding sequence and/or the deletion of some or all of the coding sequence. Gene disruptions can be generated by homologous recombination essentially as described by Rothstein (in *Methods in Enzymology*, Wu et al., eds., vol 101:202-211, 1983).

To disrupt the Och1 gene by homologous recombination, an Och1 knock-out vector can be constructed in such a way to include a selectable marker gene. The selectable marker gene is operably linked, at both 5' and 3' end, to portions of the

Och1 gene of sufficient length to mediate homologous recombination. The selectable marker can be one of any number of genes which either complement host cell auxotrophy or provide antibiotic resistance, including *URA3*, *LEU2* and *HIS3* genes. Other suitable selectable markers include the *CAT* gene, which confers chloramphenicol resistance on yeast cells, or the *lacZ* gene, which results in blue colonies due to the expression of active  $\beta$ -galactosidase. Linearized DNA fragments of an Och1 knock-out vector are then introduced into host methylotrophic yeast cells using methods well known in the art. Integration of the linear fragments into the genome and the disruption of the Och1 gene can be determined based on the selection marker and can be verified by, for example, Southern Blot analysis.

Alternatively, an Och1 knock-out vector can be constructed in such a way to include a portion of the Och1 gene to be disrupted, which portion is devoid of any Och1 promoter sequence and encodes none or an inactive fragment of the Och1 protein. By "an inactive fragment", it is meant a fragment of the Och1 protein which has, preferably, less than about 10% and most preferably, about 0% of the activity of the full-length Och1 protein. Such portion of the Och1 gene is inserted in a vector in such a way that no known promoter sequence is operably linked to the Och1 sequence, but that a stop codon and a transcription termination sequence are operably linked to the portion of the Och1 gene. This vector can be subsequently linearized in the portion of the Och1 sequence and transformed into a methylotrophic yeast strain using any of the methods known in the art. By way of single homologous recombination, this linearized vector is then integrated in the Och1 gene. Two Och1 sequences are produced in the chromosome as a result of the single homologous recombination. The first Och1 sequence is the portion of the Och1 gene from the vector, which is now under control of the Och1 promoter of the host methylotrophic yeast, yet cannot produce an active Och1 protein as such Och1 sequence codes for no or an inactive fragment of the Och1 protein, as described hereinabove. The second Och1 sequence is a full Och1 coding sequence, but is not operably linked to any known promoter sequence and thus, no active messenger is expected to be formed for

synthesis of an active Och1 protein. Preferably, an inactivating mutation is introduced in the Och1 sequence, to the 5' end of the site of linearization of the vector and to the 3' end of the translation initiation codon of Och1. By "inactivating mutation" it is meant a mutation introducing a stop codon, a frameshift mutation or any other mutation causing a disruption of the reading frame. Such mutation can be introduced into an Och1 sequence using any of the site directed mutagenesis methods known in the art. Such inactivating mutation ensures that no functional Och1 protein can be formed even if there exist some promoter sequences 5' to the Och1 sequence in the knock-out vector.

10 A preferred Och1 knock-out vector of the present invention is pBLURAS<sup>+</sup>PpOch1.

If desired, either or both of a mannosidase expression sequence and a glucosidase expression sequence can be carried on the same plasmid used to disrupt the Och1 gene to create a "knock-in-and-knock-out" vector.

15 Additionally, any of the above-described vectors can further include a nucleotide sequence capable of expressing a glycoprotein of interest in a methylotrophic yeast strain.

Another aspect of the present invention is directed to methods of modifying methylotrophic yeast strains to reduce glycosylation on proteins produced by the methylotrophic yeast strains. In accordance with the present methods, methylotrophic yeast strains are modified by transforming into these yeast strains with one or more, i.e., at least one, knock-in and/or knock-out vectors of the present invention as described herein above.

20 Methylotrophic yeast strains which can be modified using the present methods include but are not limited to yeast capable of growth on methanol such as yeasts of the genera *Candida*, *Hansenula*, *Torulopsis*, and *Pichia*. A list of species which are exemplary of this class of yeasts can be found in C. Anthony (1982), *The Biochemistry of Methylotrophs*, 269. *Pichia pastoris*, *Pichia melanolica*, *Pichia anomala*, *Hansenula polymorpha* and *Candida boidinii* are examples of

methylophilic yeasts useful in the practice of the present invention. Preferred methylophilic yeasts are of the genus *Pichia*. Especially preferred are *Pichia pastoris* strains GS115 (NRRL Y-15851); GS190 (NRRL Y-18014) disclosed in U.S. Patent No. 4,818,700; PPF1 (NRRL Y-18017) disclosed in U.S. Patent No. 4,812,405; PPY120H and yGC4; as well as strains derived therefrom.

Methylophilic yeast strains which can be modified using the present methods also include those methylophilic yeast strains which have been genetically engineered to express one or more heterologous glycoproteins of interest. The glycosylation on the heterologous glycoproteins expressed from these previously engineered strains can be reduced by transforming such strains with one or more of the vectors of the present invention.

The vectors of the present invention can be introduced into the cells of a methylophilic yeast strain using known methods such as the spheroplast technique, described by Cregg et al. 1985, or the whole-cell lithium chloride yeast transformation system, Ito et al. *Agric. Biol. Chem.* 48:341, modified for use in *Pichia* as described in EP 312,934. Other published methods useful for transformation of the plasmids or linear vectors include U.S. Patent No. 4,929,555; Himmen et al. *Proc. Nat. Acad. Sci. USA* 75:1929 (1978); Ito et al. *J. Bacteriol.* 153:163 (1983); U.S. Patent No. 4,879,231; Sreekrishna et al. *Gene* 59:115 (1987). Electroporation and PEG1000

whole cell transformation procedures may also be used. Cregg and Russel *Methods in Molecular Biology: Pichia Protocols*, Chapter 3, Humana Press, Totowa, N.J., pp. 27-39 (1998).

Transformed yeast cells can be selected by using appropriate techniques including but not limited to culturing auxotrophic cells after transformation in the absence of the biochemical product required (due to the cell's auxotrophy), selection for and detection of a new phenotype, or culturing in the presence of an antibiotic which is toxic to the yeast in the absence of a resistance gene contained in the transformants. Transformants can also be selected and/or verified by integration of the

expression cassette into the genome, which can be assessed by e.g., Southern Blot or PCR analysis.

In one embodiment, a methylophilic yeast strain is transformed with a vector which includes a nucleotide sequence coding for an  $\alpha$ -1,2-mannosidase or a functional part thereof. The nucleotide sequence is capable of expressing the  $\alpha$ -1,2-mannosidase or the functional part in the methylophilic yeast strain, and is, preferably, integrated into the genome of the methylophilic yeast strain.

The expression of an  $\alpha$ -1,2-mannosidase introduced in a methylophilic yeast strain can be verified both at the mRNA level, e.g., by Northern Blot analysis, and at the protein level, e.g., by Western Blot analysis. The intracellular localization of the protein can be analyzed by using a variety of techniques, including subcellular fractionation and immunofluorescence experiments. An ER localization of an  $\alpha$ -1,2-mannosidase can be determined by co-sedimentation of this enzyme with a known ER resident protein (e.g., Protein Disulfide Isomerase) in a subcellular fractionation experiment. An ER localization can also be determined by an immunofluorescence staining pattern characteristic of ER resident proteins, typically a perinuclear staining pattern.

To confirm that an  $\alpha$ -1,2-mannosidase or a functional part thereof expressed in a methylophilic yeast strain has the expected mannoside-trimming activity, both *in vitro* and *in vivo* assays can be employed. Typically, an *in vitro* assay involves digestion of an *in vitro* synthesized substrate, e.g., Man<sub>9</sub>GlcNAc<sub>2</sub>, with the enzyme expressed and purified from a methylophilic yeast strain, and assessing the ability of such enzyme to trim Man<sub>9</sub>GlcNAc<sub>2</sub> to, e.g., Man<sub>5</sub>GlcNAc<sub>2</sub>. In *in vivo* assays, the  $\alpha$ -1,2-mannosidase or a part thereof is co-expressed in a methylophilic yeast with a glycoprotein known to be glycosylated with N-glycans bearing terminal  $\alpha$ -1,2-linked mannose residues in such yeast. The enzymatic activity of such an  $\alpha$ -1,2-mannosidase or a part thereof can be measured based on the reduction of the number of  $\alpha$ -1,2-linked mannose residues in the structures of the N-glycans of the

glycoprotein. In both *in vitro* and *in vivo* assays, the composition of a carbohydrate group can be determined using techniques that are well known in the art and are illustrated in the Examples hereinbelow.

In another embodiment, a methylotrophic yeast strain is transformed with a vector which includes a nucleotide sequence coding for a glucosidase II or a functional part thereof. The nucleotide sequence is capable of expressing the glucosidase II or the functional part in the methylotrophic yeast strain, and is, preferably, integrated into the genome of the methylotrophic yeast strain.

The enzymatic activity of a glucosidase II or a functional part thereof expressed in a transformed methylotrophic yeast strain can be assessed using a variety of assays. For example, methylotrophic yeast cells transformed with a sequence encoding a glucosidase II or a part thereof can be set to grow on solid medium containing a substrate of the glucosidase, e.g., 5-bromo-4-chloro-3-indolyl- $\alpha$ -D-glucopyranoside or 4-MU- $\alpha$ -D-Glc. When the enzyme is expressed by the *Pichia* and secreted extracellularly, the substrate is acted upon by the enzyme, giving rise to detectable signals around the colonies such as blue color or fluorescent glow. Alternatively, liquid culture medium containing the expressed protein molecules can be collected and incubated in test tubes with a substrate, e.g., p-nitrophenyl- $\alpha$ -D-glucopyranoside. The enzymatic activity can be determined by measuring the specific product released. Moreover, *in vivo* assays can be employed, where a glucosidase II is co-expressed in yeast with a glycoprotein known to be N-glycosylated with glucose residues, e.g., influenza neuraminidase. The enzymatic activity of the glucosidase II can be measured based on the reduction of the glucose content in the sugar chain(s) of the glycoprotein.

In still another embodiment of the present invention, a methylotrophic yeast strain is transformed with an Och1 knock-out vector. As a result of the transformation and integration of the vector, the genomic Och1 gene in the yeast strains is disrupted.

In a further embodiment of the present invention, a methylotrophic yeast strain is transformed with any combination of an  $\alpha$ -1,2-mannosidase expression vector, a glucosidase II expression vector, and an Och1 knock-out vector. Such modification can be achieved by serial, consecutive transformations, i.e., introducing one vector at a time, or alternatively by co-transformation, i.e., introducing the vectors simultaneously.

The modified methylotrophic yeast strains described herein above can be further modified if desired. For example, additional disruption of genes encoding any other *Pichia* mannosyltransferases can be made. Genes encoding mammalian enzymes can also be introduced to produce glycoproteins having hybrid- or complex-type N-glycans, if desired.

Methylotrophic yeast strains which are modified by using the present methods, i.e., by transforming with one or more of the vectors of the present invention, form another embodiment of the present invention.

It should be understood that certain aspects of the present invention, especially the introduction of an intracellularly expressed  $\alpha$ -1,2-mannosidase activity, are also useful to obtain a reduced glycosylation of the O-linked glycans on glycoproteins produced in a methylotrophic yeast, as it is known in the art that these O-linked glycans consist mainly of  $\alpha$ -1,2-linked mannose residues. O-linked glycans as used herein refers to carbohydrate structures linked to serine or threonine residues of glycoproteins.

A further aspect of the invention is directed to methods of producing a glycoprotein with reduced glycosylation in a methylotrophic yeast, especially a glycoprotein heterologous to the methylotrophic yeast.

"A glycoprotein" as used herein refers to a protein which, in methylotrophic yeasts, is either glycosylated on one or more asparagines residues or on one or more serine or threonine residues, or on both asparagines and serine or threonine residues.

The term "reduced glycosylation" refers to a reduced size of the

carbohydrate moiety on the glycoprotein, particularly with fewer mannose residues, when the glycoprotein is expressed in a methylotrophic yeast strain which has been modified in accordance with the present invention, as compared to a wild type, unmodified strain of the methylotrophic yeast.

In accordance with the present invention, the production of a glycoprotein of interest with reduced glycosylation can be achieved in a number of ways. A nucleotide sequence capable of expressing a glycoprotein can be introduced into a methylotrophic yeast strain which has been previously modified in accordance with the present invention, i.e., a strain transformed with one or more of the vectors of the present invention and capable of producing glycoproteins with reduced glycosylation. Alternatively, a methylotrophic yeast strain which has already been genetically engineered to express a glycoprotein can be transformed with one or more of the vectors of the present invention. Otherwise, if a methylotrophic yeast strain does not express a glycoprotein of interest, nor is the strain transformed with any of the vectors of the present invention, such yeast strain can be transformed, either consecutively or simultaneously, with both a nucleotide sequence capable of expressing the glycoprotein and one or more vectors of the present invention. Additionally, a methylotrophic yeast strain can be transformed with one or more of the present knock-in and/or knock-out vectors which also include a nucleotide sequence capable of expressing a glycoprotein in the methylotrophic yeast strain.

The nucleotide sequence capable of expressing a glycoprotein in a methylotrophic yeast can be made to include from 5' to 3', a promoter, a sequence encoding the glycoprotein, and a 3' termination sequence. Promoters and 3' termination sequences which are suitable for expression of a glycoprotein can include any of those promoters and 3' termination sequences described hereinabove.

The nucleotide sequence for expression of a glycoprotein can include additional sequences, e.g., signal sequences coding for transit peptides when secretion of a protein product is desired. Such sequences are widely known, readily available

and include *Saccharomyces cerevisiae* alpha mating factor prepro (amf), *Pichia pastoris* acid phosphatase (PHO1) signal sequence and the like.

The nucleotide sequence for expression of a glycoprotein can be placed on a replicative vector or an integrative vector. The choice and construction of such vectors are as described hereinabove.

The nucleotide sequence capable of expressing a glycoprotein can be carried on the same replicative plasmid as a plasmid-borne  $\alpha$ -1,2-mannosidase or glucosidase II expression unit. Alternatively, the nucleotide sequence containing the glycoprotein coding sequence is carried on a separate plasmid or integrated into the host genome.

Glycoproteins produced can be purified by conventional methods.

Purification protocols can be determined by the nature of the specific protein to be purified. Such determination is within the ordinary level of skill in the art. For example, the cell culture medium is separated from the cells and the protein secreted from the cells can be isolated from the medium by routine isolation techniques such as precipitation, immunoadsorption, fractionation or a variety of chromatographic methods.

Glycoproteins which can be produced by the methods of the present invention include, e.g., *Bacillus amyloliquefaciens*  $\alpha$ -amylase, *S. cerevisiae* invertase,

*Trypanosoma cruzi* trans-sialidase, HIV envelope protein, influenza virus A haemagglutinin, influenza neuraminidase, Bovine herpes virus type-1 glycoprotein D, human angiotensin, human B7-1, B7-2 and B-7 receptor CTLA-4, human tissue factor, growth factors (e.g., platelet-derived growth factor), tissue plasminogen activator, plasminogen activator inhibitor-1, urokinase, human lysosomal proteins such as  $\alpha$ -

galactosidase, plasminogen, thrombin, factor XIII and immunoglobulins. For additional useful glycoproteins which can be expressed in the genetically engineered *Pichia* strains of the present invention, see Brethauer and Castellino, *Biotechnol.*

*Appl. Biochem.* 30: 193-200 (1999), and Kukuruzinska et al. *Ann Rev. Biochem.* 56: 915-44 (1987).

Glycoproteins produced by using the methods of the present invention, i.e., glycoproteins with reduced glycosylation, are also part of the present invention.

5 Still another aspect of the present invention provides kits which contain one or more of the knock-in vectors, knock-out vectors, or knock-in-and-knock-out vectors of the present invention described above. More particularly, a kit of the present invention contains a vector capable of expressing an  $\alpha$ -mannosidase I in a methylotrophic yeast, a vector capable of expressing a glucosidase II in a

10 methylotrophic yeast, a vector capable of disrupting the *Och1* gene in a methylotrophic yeast, a vector capable of expressing both a glucosidase II and an  $\alpha$ -mannosidase, a vector a vector capable of disrupting the *Och1* gene and capable of expressing either or both of a glucosidase II and an  $\alpha$ -mannosidase, or any combinations thereof.

15 The kit can also include a nucleotide sequence which encodes and is capable of expressing a heterologous glycoprotein of interest. Such nucleotide sequence can be provided in a separate vector or in the same vector which contains sequences for knocking-in or knocking out as described hereinabove.

In addition, the kit can include a plasmid vector in which a nucleotide sequence encoding a heterologous protein of interest can be subsequently inserted for transformation into and expression in a methylotrophic yeast. Alternatively, the knock-in or knock-out vectors in the kits have convenient cloning sites for insertion of a nucleotide sequence encoding a heterologous protein of interest.

The kit can also include a methylotrophic yeast strain which can be subsequently transformed with any of the knock-in, knock-out or knock-in-and-knock-out vectors described hereinabove. The kit can also include a methylotrophic yeast strain which has been transformed with one or more of the knock-in or knock-out vectors. Furthermore, the kit can include a methylotrophic yeast strain which has been

transformed with a nucleotide sequence encoding and capable of expressing a heterologous glycoprotein of interest.

The present invention is further illustrated by the following examples.

Example 1

Introduction of  $\alpha$ -1,2-Mannosidase to the ER-Golgi Border

1.1 Plasmids

Plasmid	Promoter	Enzyme	Tag
pGAPZMFManHDEL	GAP	<i>T. reesei</i> $\alpha$ -1,2-mannosidase	-
pGAPZMFManMycHDEL	GAP	<i>T. reesei</i> $\alpha$ -1,2-mannosidase	Myc
pPICZBMFManMycHDEL	AOXI	<i>T. reesei</i> $\alpha$ -1,2-mannosidase	Myc
pGAPZMFmManHDEL	GAP	mouse mannosidase IB	-
pGAPZMFmMycManHDEL	GAP	mouse mannosidase IB catalytic domain	Myc

The *Trichoderma reesei*  $\alpha$ -1,2-mannosidase gene has been isolated and described by Maras et al. (*J. Biotechnol.* 77:255-263, 2000). The sequence of this gene is available at NCBI Genbank under Accession No. AF212153. A construction fragment was generated by PCR using the pPIC9Mfmanase plasmid (same as pPP1MFmds1 described by Maras et al. (2000)) as the template and using the following oligonucleotide primers: 5'-GACTGGTTCCAAATTGACAAAGC-3' (SEQ ID NO:2) and 5'-AGTCTAGATTACAACTCGTCGTGAGCAAGGTGCCGCCCGGTCG-3' (SEQ ID NO:3). The resulting product contained the 3' end of the *Pichia pastoris* AOXI promoter, the prepro-signal sequence of the *S. cerevisiae*  $\alpha$ -mating factor, the open reading frame of the *Trichoderma reesei*  $\alpha$ -1,2-mannosidase cloned in frame with the signal sequence, the coding sequence for HDEL, a stop codon and an *Xba* I restriction site. This fragment was digested with *Eco* RI and *Xba* I, removing the 5' sequences up to the mannosidase ORF, and then cloned into the vector pGAPZaA (Invitrogen, Baarn, The Netherlands) which had been digested with *Eco* RI and *Xba* I, thus restoring the fusion with the *S. cerevisiae*  $\alpha$ -mating factor signal

sequence. The resulting plasmid was named pGAPZMFManHDEL and is graphically depicted in Figure 1. The ORF sequence of the MFManHDEL fusion in pGAPZMFManHDEL is set forth in SEQ ID NO: 14.

In order to introduce the coding sequence for a c-Myc tag between the catalytic domain and the HDEL-signal, the 3' end of the ORF of *T. reesei*  $\alpha$ -1,2-mannosidase was PCR-amplified using a sense primer 5'-

CCATTGAGGACGCATGCCCGCC-3' (SEQ ID NO: 4) (containing an *Sph* I restriction site) and an antisense primer

GTATCTAGATTACAACCTCGTCGTGCAGATCCTCTTCTGAGATGAGTTTGT

TCAGCAAGGTGGCCGCCCGTCGTGATGAA (SEQ ID NO: 5) (containing the coding sequences of the c-Myc tag and the HDEL signal, followed by a stop codon and an *Xba* I restriction site). The resulting PCR product was digested with *Sph* I and *Xba* I, purified by agarose gel electrophoresis and inserted into pGAPZMFManHDEL which had been cut with the same restriction enzymes, resulting in plasmid

pGAPZMFManMycHDEL. To put the ORF of pGAPZMFManMycHDEL under the control of the inducible AOXI promoter, the entire ORF was liberated from pGAPZMFManMycHDEL with *Bst* BI and *Xba* I, and cloned in pPICZB (Invitrogen, Baarn, The Netherlands), resulting in pPICZBMFManMycHDEL.

Cloning of the mouse mannosidase IB catalytic domain with concomitant addition of the coding sequence for a C-terminal HDEL-tag was done by PCR on a mouse cDNA library (mRNA isolated from the L929 cell line induced with cycloheximide and mouse Tumor Necrosis Factor. Average insert length of the cDNA library was 2000 bp). The PCR oligonucleotide primers used were:

5'AACCTCGAGATGGACCTCTTCAAAACACAAAGGC3' (SEQ ID NO: 6) and

5'TTCGGCGCGCTTACAACCTCGTCGTGGACACGACGAGATTACCTGA3' (SEQ ID NO: 7). The product contained a 5' *Xho* I site and the coding sequence for C-terminal HDEL-site, followed by a stop codon and a *Not* I site at the 3' end. The product was cloned in pGAPZaA via the *Xho* I/*Not* I sites in the PCR product and the vector, resulting in an in frame fusion of the mouse mannosidase catalytic domain

with the *S. cerevisiae*  $\alpha$ -mating factor signal sequence. The sequence of the entire open reading frame generated is set forth in SEQ ID NO: 15.

## 1.2 Yeast Transformation and Genomic Integration

5

Table 2

Parental strain	DNA transformed
GS115 (his4)	pGAPZMFManHDEL
	pPIC9MFMManHDEL
	pPIC9mManHDEL
	pPIC9mMycManHDEL
	pGAPZmManHDEL
	pGAPZmMycManHDEL
GS115 (his4 complemented by pPIC9InfluenzaHA)	pGAPZMFManHDEL
	pGAPZmManHDEL
	pGAPZmMycManHDEL
PPY120H (his4 complemented by pPIC9sOCHI)	pGAPZMFManMycHDEL
	pPICZBMFMManMycHDEL
	pPIC9InfluenzaNeuraminidase
yGC4 (his4 arg1 ade2 ura3 complemented by pBLURA5'PpOCHI)	pGAPZMFManHDEL
	pPIC9Glucosylase

All transformations to *Pichia pastoris* were performed with electroporation according to the directions of Invitrogen. Transformants of vectors carrying the Zeocin resistance gene were selected on YPD containing 100  $\mu$ g/ml Zeocine

(Invitrogen, Baarn, the Netherlands) and 1M sorbitol. Selection of transformants of pPIC9 derivatives was done on minimal medium lacking histidine and containing 1M sorbitol. Genomic integration of the expression cassettes was verified using PCR on genomic DNA purified from the *Pichia* strains using the Yeast Miniprep method (Nucleon). In all cases concerning the *Trichoderma reesei* gene fusions, the primers used were the sense primer 5'-CCATTGAGGACGCGATGCGCGCC-3' (SEQ ID NO: 8), which annealed to the 3' half of the mannosidase ORF, and the antisense primer 3' AOXI 5'-GCAAAATGGCATTCTGACATCCT-3' (SEQ ID NO: 9), which annealed to the AOXI transcription terminator that was present in all our expression constructs.

10 For the control of genomic integration of the mouse mannosidase transgenes, PCR was done using the sense primer 5'GAP 5'GTCCCTATTTCATCAATCAATTGAA3' (SEQ ID NO: 10, annealing to the GAP promoter or 5'AOXI 5'GACTGGTTCCTCAATTGACAAAGC3' (SEQ ID NO:11), annealing to AOXI promoter), and the antisense primer 3'AOXI (above). For the expression constructs containing a Myc tagged *Trichoderma reesei*  $\alpha$ -1,2-mannosidase expression unit, further evidence for genomic integration was obtained using Southern Blotting with the entire MFManMycHDEL ORF (<sup>32</sup>P labelled using HighPrime, Boehringer Mannheim) as a probe.

20

### 1.3 Expression of $\alpha$ -1,2-mannosidase

Expression of an  $\alpha$ -1,2-Mannosidase in GS115 strains expressing influenza virus haemagglutinin was verified by qualitative Northern blot. Expression of an  $\alpha$ -1,2-Mannosidase in PPY120H strains was verified by anti-Myc Western blot.

25

*Qualitative Northern Blot* – Total RNA was purified from *Pichia* strains and the yield was determined spectrophotometrically. Northern blotting was performed according to standard procedures and an estimate of the quantity of RNA loaded was made using methylene blue staining of the blot, visualizing the rRNA

bands. The blot was probed with a *ClaI/NarI* fragment of the mannosidase, labelled with <sup>32</sup>P using HighPrime (Boehringer Mannheim).

*SDS-PAGE and Western Blotting* – Total yeast cell lysates were prepared by washing the cells twice with PBS, followed by boiling in 1 volume of 2x concentrated Laemmli loading buffer for 5 min. The lysate was spun briefly in a microcentrifuge prior to gel loading and only the supernatant was loaded. For the analysis of proteins secreted into the growth media, the proteins were precipitated from 200 µl of these media using desoxycholate/trichloroacetic acid according to standard procedures. The pellet was redissolved in 2x concentrated Laemmli loading buffer and the solutions were pH-corrected using Tris. SDS-PAGE was performed and followed by semidry electroblotting to nitrocellulose membranes. For Western Blotting, the 9E10 anti-Myc and the anti-HA mouse monoclonals (Boehringer Mannheim) were used at a concentration of 1 µg/ml, and the rabbit anti-PDI antiserum (Stressgen) was used at a dilution of 1/500. The secondary antibodies were goat anti-mouse IgG conjugated to alkaline phosphatase for the monoclonals and goat anti-rabbit IgG conjugated to peroxidase for the polyclonal (secondary antibodies from Sigma). Detection was performed using the NBT/BCIP system for alkaline phosphatase and the Renaissance substrate (NEN/Biosciences) for peroxidase. Imaging of the latter blot result was done on a Lumilager imaging device (Boehringer Mannheim).

The results shown in Figure 4 indicated that the great majority of the HDEL-tagged protein was retained intracellularly, both when expressed from the strong constitutive GAP promoter and when expressed from the strong inducible AOXI promoter.

#### 1.4 Localization of $\alpha$ -1,2-Mannosidase

*Isopycnic sucrose density gradient centrifugation* – To determine the localization of the HDEL-tagged mannosidase, subcellular fractionation was carried

out using cells expressing the mannosidase-Myc-HDEL from the strong constitutive GAP promoter.

Briefly, 0.5 g of wet weight yeast cells were lysed using 4 x 1 ml vortexing with 4.5 g glass beads in 1 ml lysis-buffer (50 mM Tris-HCl pH 7.5 containing 0.6 M sorbitol, 10 mM  $\beta$ -mercaptoethanol and 5 mM MgCl<sub>2</sub>). Between vortexing periods, the mixture was placed on ice for 5 min. The supernatant was collected and the glass beads were washed once with lysis-buffer, and the supernatant of this washing step was added to the first supernatant. This lysate was subjected to a differential centrifugation procedure. The P10000 pellet was solubilized in 0.5 ml of a 60% sucrose solution in lysis buffer. This solution was placed at the bottom of an Ultraclear ultracentrifuge tube (Beckman) of 14 x 89 mm. Subsequently, 1.5 ml each of sucrose solutions of 55, 50, 45, 42.5, 40, and 37.5% were carefully layered over each other. The tube was filled to the edge with 35% sucrose. Isopycnic sucrose gradient centrifugation was performed for 14 h at 180,000g in a Beckman SW 41 rotor in a Beckman Model L8-70 preparative ultracentrifuge. After completion, 1ml fractions were collected from the top and partially dialysed from excess sucrose, evaporated to dryness in a vacuum centrifuge. After redissolving the pellet in Laemmli buffer, the samples were subjected to SDS-PAGE in triplicate and the Western blots were treated with anti-HA, anti-Myc or anti-PDI ("PDI" for Protein Disulfide Isomerase), respectively.

The results illustrated almost exact cosedimentation of the MFManMycHDEL protein with the Protein Disulfide Isomerase marker protein (which is also targeted with a HDEL signal sequence) (Figure 5). In the same assay, the HA-tagged OCHI was distributed over the whole gradient, with the highest abundance in fractions having a density lower than that of the fractions containing the mannosidase and the PDI. This result indicated that the mannosidase was targeted to the expected location (the ER-Golgi boundary) by the addition of an HDEL signal. In contrast, the mannosidase without HDEL, expressed from inducible alcohol oxidase I

promoter (which was of comparable strength as the GAP promoter), was secreted at a high level of about 20 mg/l.

*Immunofluorescence microscopy* – To confirm the correct targeting of the mannosidase-Myc-HDEL, an immunofluorescence microscopy experiment was performed.

Briefly, yeast cultures were grown to OD<sub>600</sub> in YPD (for pGAPZMFManMycHDEL) or in YMP following a YPGlycerol growth phase for pPICZBMFManMycHDEL. Formaldehyde was added to the yeast cultures to a final concentration of 4% and incubated for 10 min at room temperature. Cells were pelleted and resuspended in 50mM potassium phosphate buffer pH 6.5 containing 1mM MgCl<sub>2</sub> and 4% formaldehyde and incubated for 2h at room temperature. After pelleting, the cells were resuspended to an OD<sub>600</sub>=10 in 100 mM potassium phosphate buffer pH 7.5 containing 1mM MgCl<sub>2</sub> and EDTA-free Complete™ protease inhibitor cocktail (Boehringer Mannheim). To 100 µl of cell suspension, 0.6 µl of β-mercaptoethanol and 20µl of 20,000 U/ml Zymolyase 100T (ICN) were added, followed by a 25 minute incubation with gentle shaking. The cells were washed twice in the incubation buffer and added to poly-lysine coated cover slips (these are prepared using adhesive rings normally in use for reinforcing perforations in paper). Excess liquid was blotted with a cotton swab and the cells were allowed to dry at 20°C. All blocking, antibody incubation and washing steps are performed in PBS containing 0.05% bovine serum albumin. Primary antibodies are used at 2µg/µl and secondary antibodies conjugated to fluorophores (Molecular probes) were used at 5µg/µl. The nucleus was stained with the nucleic acid stain HOECHST 33258. After fixation and cell wall permeabilization, the integrity of the yeast cell morphology was checked in phase contrast microscopy and after immunostaining, the slides were examined under a Zeiss Axiophot fluorescence microscope equipped with a Kodak digital camera. Images were processed using Macprobe 4.0 software and prepared with Corel Photopaint 9.0.

The Golgi marker protein OCH1-HA gave the typical Golgi staining pattern described in the literature (speckle-like staining). Staining with the 9E10 monoclonal anti-Myc antibody, recognizing mannosidase-Myc-HDEL, gave a perinuclear staining pattern with some disparate staining in the cytoplasm, highly indicative for an ER targeting (Figure 4).

Based on the foregoing experiments, it is concluded that the *Trichoderma reesei* mannosidase-Myc-HDEL was targeted to the ER-Golgi boundary.

## Example 2

### Co-expression of Mannosidase-HDEL with Recombinant Glycoproteins

#### Co-expression of Mannosidase-HDEL with the *Trypanosoma cruzi* *trans*-Sialidase

The cloning of a *Trypanosoma cruzi* *trans*-sialidase gene coding for an active *trans*-sialidase member without the C-terminal repeat domain has been described by Laroy et al. (*Protein Expression and Purification* 20: 389, 2000) which is incorporated herein by reference. The sequence of this *Trypanosoma cruzi* *trans*-sialidase gene is available through NCBI Genbank under the Accession No. AJ276679. For expression in *P. pastoris*, the entire gene was cloned in pHILD2 (Invitrogen, San Diego, CA), creating pHILD2-TS. To allow better secretion, pPIC9-TS was created in which *trans*-sialidase was linked to the prepro secretion signal of the yeast α-mating factor. Plasmids pPIC9-TSE and pCAGGS-prepro-TSE were created where the epitope E-tag was added to the C-terminal of the *trans*-sialidase to allow easy detection and purification. The construction of pHILD2-TS, pPIC9-TSE and pCAGGS-prepro-TSE has been described by Laroy et al. (2000), incorporated herein by reference. The vectors used in the construction were made available through <http://www.belspo.be/bccm/lmbp.htm#main> for pCAGGS (No. LMBP 2453), Invitrogen, San Diego, CA for pHILD2 and pPIC9, and Pharmacia Biotech for pCANTAB-5E.

Plasmid pPIC9-TSE was linearized with SstI and was transformed into *P. pastoris* GS115 (*His4*) strain by electroporation according to the manufacturer's instructions (Invitrogen). One of the transformants was further transformed with plasmid pGAPZMFManHDEL, establishing a strain co-expressing Mannosidase-HDEL and the *Trypanosoma cruzi trans-sialidase*.

Fermentation and protein purification was according to the procedures described by Laroy et al. (2000).

Purified *trans*-sialidase was subject to carbohydrate analysis according to Callewaert et al., *Glycobiology* 11, 4, 275-281, 2001. Briefly, the glycoproteins were bound to the PVDF membrane in the wells of a 96-well plate, reduced, alkylated and submitted to peptide-N-glycosidase F deglycosylation. The glycans were derivatised with 8-amino-1,3,6-pyrenetrisulfonic acid by reductive amination. Subsequently, the excess free label was removed using Sephadex G10-packed spin columns and the glycans were analysed by electrophoresis on a 36 cm sequencing gel on an ABI 377A DNA-sequencer and detected using the built-in argon laser. Digests with 3 mU/ml purified *T. reesei*  $\alpha$ -1,2-mannosidase (described by Maras et al., *J. Biotechnol.* 77, 255-63, 2000) were also performed in 20 mM sodium acetate pH=5.0. The glycans derived from 1  $\mu$ g of the purified recombinant glycoproteins were used as the substrate. IU of the  $\alpha$ -1,2-mannosidase is defined as the amount of enzyme that releases 1  $\mu$ mol of mannose from baker's yeast mannan per minute at 37°C and pH=5.0.

As can be seen in Figure 6, panel B, the major N-glycan on *trans*-sialidase was Man<sub>5</sub>GlcNAc<sub>2</sub> (Compare with panel F, representing an analysis of the N-glycans of bovine RNaseB. The one but last peak in this profile is Man<sub>5</sub>GlcNAc<sub>2</sub>, the first peak is Man<sub>5</sub>GlcNAc<sub>6</sub>). *In vitro*, this glycan was digestible to Man<sub>5</sub>GlcNAc<sub>2</sub> with  $\alpha$ -1,2-mannosidase (Figure 6, panel C). In the N-glycan profile of the *trans*-sialidase co-expressed with mannosidase-HDEL, the major peak corresponded to Man<sub>5</sub>GlcNAc<sub>2</sub> (Figure 6, panel D).

#### Co-expression of Mannosidase-HDEL with the Influenza A virus haemagglutinin

The Influenza A virus haemagglutinin was known to be glycosylated in *Pichia pastoris* with high-mannose N-glycans containing 9-12 mannose residues (Saelens et al. *Eur. J. Biochem.* 260: 166-175, 1999). The effect of a co-expressed mannosidase on the N-glycans of the haemagglutinin was assessed in an N-glycan profiling method described below. In addition, to compare the efficiency of the *Trichoderma* enzyme (having a temperature optimum of 60°C) with a mammalian mannosidase having a temperature optimum of 37°C, the catalytic domain of the mouse mannosidase IB from a mouse cDNA-library was cloned and tagged with a HDEL signal by PCR amplification. This ORF was cloned after the prepro-signal sequence of the *S. cerevisiae*  $\alpha$ -mating factor under the control of the GAP promoter. Expression of the mannosidase-HDEL transgenes on the mRNA level was confirmed by qualitative Northern blotting.

The haemagglutinin was expressed and purified from a non-mannosidase expressing control strain and from a strains co-expressing the *Trichoderma reesei* mannosidase-HDEL or the mouse mannosidase IB-HDEL according to the procedure described by Kulakosky et al. *Glycobiology* 8: 741-745 (1998). The purified haemagglutinin was subjected to PNGase F digestion as described by Saelens et al. *Eur. J. Biochem.* 260: 166-175, 1999. The proteins and glycans were precipitated with 3 volumes of ice-cold acetone and the glycans were extracted from the pellet with 60% methanol. Following vacuum evaporation, the glycans were labeled with 8-amino-1,3,6 pyrenetrisulfonic acid by adding 1  $\mu$ l of a 1:1 mixture of 20 mM APTS in 1.2M citric acid and 1M N<sub>4</sub>CNBH<sub>3</sub> in DMSO and incubating for 16h at 37°C at the bottom of a 250  $\mu$ l PCR-tube. The reaction was stopped by the addition of 10  $\mu$ l deionized water and the mixture was loaded on a 1.2 cm Sephadex G10 bed packed to dryness in a microspin-column by centrifugation in a swinging bucket rotor, which provided for a flat resin surface. After loading, 50  $\mu$ l deionised water was carefully added to the resin bed and the spin column was briefly centrifuged for 5 seconds at 750g in a tabletop centrifuge. This elution process was repeated twice and all the eluates were

pooled and evaporated to dryness in a Speedvac vacuum centrifuge (Savant). The labeled glycans were reconstituted in 1.5  $\mu$ l gel loading buffer containing 50% formamide and 0.5  $\mu$ l Genescan 500™, labeled with rhodamine (Perkin Elmer Bioscience), serving as an internal reference standard. This mixture was loaded on a DNA-sequencing gel containing 10% of a 19:1 mixture of acrylamide:bisacrylamide (Biorad, Hercules, CA, USA) and made up in the standard DNA-sequencing buffer (89 mM Tris, 89 mM borate, 2.2 mM EDTA). Polymerization of the gel was catalyzed by the addition of 200  $\mu$ l 10% ammoniumpersulfate solution in water and 20  $\mu$ l TEMED. The gel was of the standard 36 cm well-to-read length and was run on an Applied Biosystems Model 373A DNA-sequencing apparatus. Pre-running of the gel was done at 1000 V for 15 min. and after loading, the gel was electrophoresed for 8h at 1250 V without heating. This methodology gives a limit of detection of 10 fmol per peak. The data were analysed with Genescan 3.0 software.

As shown in Figure 7, the *Trichoderma reesei*  $\alpha$ -1,2-mannosidase

provided the most complete reduction in the number of  $\alpha$ -1,2-mannoses present on the N-glycans. The N-glycan processing by mouse mannosidase IB-HDEL was less efficient than by the *Trichoderma reesei*  $\alpha$ -1,2-mannosidase.

Despite the efficient removal of  $\alpha$ -1,2-mannoses from the N-glycans of haemagglutinin, no Man<sub>5</sub>GlcNAc<sub>2</sub> was obtained. Even after digestion of the N-glycans with 3 mU of purified *Trichoderma reesei*  $\alpha$ -1,2-mannosidase, only Man<sub>5</sub>GlcNAc<sub>2</sub> was obtained as the smallest sugar chain. These results indicated that the remaining residues were possibly  $\alpha$ -1,6-linked mannoses, originating from the initiating OCH1  $\alpha$ -1,6-mannosyltransferase enzymatic activities. OCH1 was observed to be localized to very early part of the Golgi apparatus and could act on the N-glycans of haemagglutinin before complete digestion of the Man<sub>5</sub>GlcNAc<sub>2</sub> precursor to Man<sub>5</sub>GlcNAc<sub>2</sub> by the mannosidases-HDEL. Thus, for proteins whose glycans are efficiently modified by the  $\alpha$ -1,6-mannosyltransferase, an inactivation of the OCH1

gene coding for the transferase would be desirable in order to obtain proteins with Man<sub>5</sub>GlcNAc<sub>2</sub>.

### Example 3

#### Inactivation of the *Pichia* Och1 Gene:

A *Pichia pastoris* sequence was found in the GenBank under Accession No. E12456 and was described in Japanese Patent Application No. 07145005, incorporated herein by reference. This sequence shows all typical features of an  $\alpha$ -1,6-mannosyltransferase and is most homologous to the *S. cerevisiae* OCH1, thus referred to herein as the *Pichia pastoris* Och1 gene.

First, the full ORF of the *Pichia pastoris* Och1 gene was PCR cloned in pUC18 to obtain plasmid pUC18pOch1. pUC18pOch1 was cut with HindIII, blunt-ended with T4 polymerase, then cut with XbaI, releasing a fragment containing the 5' part of the *Pichia pastoris* Och1 gene. This fragment was ligated into the vector pBLURA IX (available from the Keck Graduate Institute, Dr. James Cregg, <http://www.kgi.edu/html/noncore/faculty/cregg/cregg.htm>), which had been cut with *Eco* RI, blunt-ended with T4 polymerase, and then cut with *Nhe* I. This ligation generated pBLURA5'PpOCH1, as shown in Figure 8.

Disruption of this *Pichia* OCH1 gene in the *Pichia* genome was achieved by single homologous recombination using pBLURA5'PpOCH1, as illustrated in Figure 9. As a result of the single homologous recombination, the Och1 gene on the *Pichia* chromosome was replaced with two Och1 sequences: one consisted only about the first one third of the full Och1 ORF, the other had a full Och1 ORF without a Och1 promoter. Single homologous recombination was achieved as follows. Cells of the *Pichia* strain yGCA were transformed by electroporation with pBLURA5'PpOCH1 which had been linearized with the single cutter *Bst* BI. About 500 transformants were obtained on minimal medium containing 1M sorbitol, biotin, arginine, adenine and histidine and incubation at 27°C. Thirty-two of these transformants were picked and re-selected under the same conditions. Twelve clones were further analyzed for correct genomic integration of the cassette by PCR. Seven of the twelve URA prototrophic clones contained the cassette in the correct locus.

One of the Och1-inactivated clones was also further transformed with pGAPZMFManHDEL to produce "supertransformants". Both the Och1-inactivated clone and three supertransformants also expressing the ManHDEL were evaluated in cell wall glycan analysis as follows. Yeast cells were grown in 10 ml YPD to an OD<sub>600</sub>=2 and mannoproteins were prepared by autoclaving the yeast cells in 20 mM sodium citrate buffer pH7 for 90 min at 120 °C and recovery of the supernatant after centrifugation. Proteins were precipitated from this supernatant with 3 volumes of cold methanol. The protein preparation obtained in this way was used for N-glycan analysis using DSA-FACE as described by Callewaert et al. (2001) *Glycobiology* 11, 275-281. As shown in Figure 10, there was an increased amount of Man<sub>5</sub>GlcNAc<sub>2</sub> glycan in the Och1-inactivated clone as compared to parent strain yGCA, indicative of a reduced activity of the Och1 enzyme. In all three supertransformants which also expressed the HDEL-tagged  $\alpha$ -1,2 mannosidase, the production of Man<sub>5</sub>GlcNAc<sub>2</sub> was observed. Furthermore, upon digestion of the same glycan mixtures with 3 mU/ml purified recombinant *Trichoderma reesei*  $\alpha$ -1,2-mannosidase, more Man<sub>5</sub>GlcNAc<sub>2</sub> was formed in the strain transformed with pBLURA5'PpOCH1 than in the parent strain (Figure 11, compare panel 2 and 3).

These results confirmed that the lack of a production of Man<sub>5</sub> glycans on recombinantly produced proteins such as haemagglutinin from cells expressing  $\alpha$ -1,2-mannosidase were due to the activity of the Och1 protein. These results further indicate that the production of glycoproteins with Man<sub>5</sub> glycans could be facilitated by the inactivation of the Och1 gene.

#### Example 4

##### Expression of Glucosidase II in *Pichia pastoris*

#### 4.1 Amplification of the GLSII alpha subunit ORF from *S. cerevisiae*.

Genomic DNA was prepared from the *S. cerevisiae* strain INVS ( $\alpha$ , leu2-3, 112 his3.1, trp1-289, ura3-52), using the Nucleon kit (Amersham). A touch-down PCR reaction was performed using this genomic DNA as template and the LA TaKaRa polymerase (ImTec Diagnostics). The sequence of the PCR primers was based on the known sequence of the *S. cerevisiae* GLSII ORF:

10 Sense primer: 5' CCG CTC GAG ATG GTC CTT TTG AAA TGG CTC 3'

(SEQ ID NO:12)

Antisense primer: 5' CCG GGC CCA AAA ATA ACT TCC CAA TCT TCA

15 G 3' (SEQ ID NO:13)

#### 4.2 Cloning of the *S. cerevisiae* glucosidase II ORF into *Pichia pastoris* expression vectors.

Construction of the glucosidase II expression vectors -- The PCR fragment was digested with *Xho* I/Apo I and ligated into the pGAPZA vector (Invitrogen), thereby placing the ORF under the transcriptional control of the GAP promoter.

Using this strategy, the myc and the His6 tag were placed in frame to the C-terminus of Glucosidase II, creating pGAPZAGLSII. The complete ORF of pGAPZAGLSII was then sequenced to ensure that no mutations were generated in the PCR reaction.

25 The sequence of the vector pGAPZAGLSII was set forth in SEQ ID NO: 18. The GLSII ORF from the pGAPZAGLSII vector was cloned into vector pPICZA

(Invitrogen) to create pPICZAGLSII, thereby placing the ORF under the transcriptional control of the AOX1 promoter. The GLSII ORF from the

30 pGAPZAGLSII vector was cloned into vector pAOX2ZA, thereby placing the ORF under the transcriptional control of the AOX2 promoter. This vector was created by

replacing the multi cloning site of vector pAOX2ZB with the multi cloning site of pPICZA. Vector pAOX2ZB was generated by replacing the AOX1 promoter of pPICZB by the AOX2 promoter region of the AOX2 gene (Martinet et al., Biotechnology Letters 21). The AOX2 promoter region was generated by PCR on

5 *Pichia* genomic DNA with the sense primer

5'GACGAGATCTTTTTCAGACCATATGACCGG 3' (SEQ ID NO: 26) and the antisense primer 5'GCGGAAATCTTTTCTCAGTTGATTGTTGT 3' (SEQ ID NO: 27). The GLSII ORF from the pGAPZAGLSII vector was cloned into vector pYPT1ZA to create pYPT1ZAGLSII, thereby placing the ORF under the

10 transcriptional control of the YPT1 promoter. Vector pYPTZA was created by

replacing the AOX1 promoter of pPICZA by the YPT1 promoter present on the plasmid pIB3 (GenBank accession number AF027960)(Sears et al., Yeast 14, pg 783-790, 1998). All constructs contain the phleomycin resistance gene. The resulting final expression vectors (pGAPZAGLSII, pAOX2ZAGLSII, pPICZAGLSII and pYPT1ZAGLSII) are depicted in Figures 12-15.

15 Similar expression vectors were constructed, carrying the Ampicillin resistance marker and the *Pichia* ADE1 selection marker. In principle, the Zeocin resistance expression cassette of the plasmids pAOX2ZAGLSII, pGAPZAGLSII and pYPT1ZAGLSII was replaced by the Ampicillin and *Pichia* ADE1 cassette of the vector pBLADE IX (Cregg, J.M.) to result in the vectors pAOX2ADE1glsII, pGAPADE1glsII and pYPT1ADE1glsII. Vector pPICADE1glsII was obtained by inserting the glucosidase II open reading frame into the multiple cloning site of the vector pBLADE IX (Cregg, J.M.). The resulting final expression vectors (pGAPADE1glsII, pAOX2ADE1glsII, pPICADE1glsII and pYPT1ADE1glsII) are depicted in Figures 16-20.

20 Adding the ER retention tag HDEL to Glucosidase II expression vectors -- The following primers were used to generate an HDEL-containing PCR fragment:

Primer 1: 5'GCG GGT CGA CCA C/GA C/GA A/CT G/TG A/GT TTT AGC CTT

Sel I H D E L stop



H1-3, which indicate that the HDEL tag, which was added C-terminally to the GLSII ORF, resulted in an ER localization of the protein, preventing GLSII to be secreted into the growth medium.

*Intracellular expression* – The 6 transformants and the WT strain were grown for 2 days in 500 ml BMDY. The cells were collected by centrifugation, washed, resuspended into a minimal volume (50 mM Tris.HCl pH 7.5, 5% glycerol) and broken using glass beads. The cell debris was removed through several centrifugation steps (low speed centrifugation (2000-3000g)). Membranes were obtained from the supernatant through ultracentrifugation. The pellets were resuspended in Laemmli buffer and loaded for SDS-PAGE analysis. The proteins were blotted on a nitrocellulose membrane. The intracellular GLSII expression was checked using anti-myc Ab and peroxidase conjugated secondary Ab. Following the luminescence detection, a band at about 110 kDa was observed with the GLSIHDEL transformants (H1 and H3, faint signal for H2), but not with the WT and GLSII expression strains. These results clearly indicate the intracellular presence of the recombinant GLSII when expressed with a C-terminal HDEL tag. No GLSII was detected intracellularly when this tag was not present.

#### 4.5 Purification and activity assays of the recombinant glucosidase II alpha

##### submit

A GLSII assay was set up as follows and was tested using a commercially available yeast alpha-glucosidase (Sigma) as a positive control.

Composition: 70 µl 80 mM phosphate-citrate buffer pH 6.8, 7 µl 250 mM mannose, 3.5 µl 250 mM 2-deoxy-D-glucose, 0.8 µl 4-MeUmbelliferyl-alpha-D-glucopyranoside (1 µM). Three assays were performed: one with 1 unit commercial enzyme, one without the enzyme and one with the enzyme but without the substrate. The assay mixture was incubated overnight at 30°C. When illuminated with UV, only the reaction mixture with both the enzyme and the substrate showed fluorescence

(Figure 22). This indicates that the assay was very specific in detecting the activity of the alpha-glucosidase.

WT PPY12-OH<sub>2</sub> strain 18 and strain H3 were grown during 2 days in 2x10 ml growth medium. Cells were spun down and medium was adjusted to 300 mM NaCl and 10 mM imidazol and concentrated with VivaSpin columns to 0.5-1ml. Medium was loaded onto a Ni-NTA spin column (Qiagen) and the purification was performed according to the manufactures recommendations. Protein was eluted from the column in 2x100 µl elution buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 250 mM imidazol pH 8.0). From each eluate, 20 µl was assayed for its glucosidase II activity. 0.33 units of the commercial enzyme diluted in 20 µl of the elution buffer was used as a positive control. Fluorescence was observed with the positive control and the elute of strain 18, the strain which secreted the enzyme into the growth medium. These results indicate that the recombinant *S. cerevisiae* GLSII alpha subunit, secreted by *Pichia pastoris*, was a functionally active enzyme. The activity was not seen in the WT (untransformed) strain, nor in strain H3 as the GLSII was expressed intracellularly (Figure 23). These results also indicate that the beta subunit is not necessary for the functionality of the alpha part of the protein.

Example 5  
Creating *Pichia* Strains Expressing  
both Glucosidase II and Mannosidase

5 Strain GS115 was transformed with pGAPZGLSII and pGAPZglsIIHDEL. Transformants were selected on YPDSzeo.

Strain yGC4 was transformed with the following constructs, respectively:

(1) pGAPADEglsII and pGAPADEglsIIHDEL, selection on synthetic sorbitol medium without adenine;

(2) pGAPZMFManHDEL: selection on YPDSzeo; and

(3) pGAPZMFManHDEL/pGAPADEglsIIHDEL: selection on synthetic sorbitol medium without adenine and with zeocin.

15 Strain yGC4 with OCH1 knock-in and expressing MfmannosidaseHDEL was transformed with pGAPADEglsII and pGAPADEglsIIHDEL. Selection of transformants was done on synthetic sorbitol medium without adenine and uracil.

For all transformations, colonies were obtained. Transformants with the expression vector(s) integrated into the genome, determined by PCR, were obtained. Expression of GLSII from some of these transformants was observed.

SEQUENCE LISTING

SEQ ID NO: 1  
HDEL (peptide)

5 5'-GACTGGTTCCAAATTGACAAAGC-3'

SEQ ID NO: 2

SEQ ID NO: 3

5'-AGTCTAGATTACAACCTCGTCGTGAGCAAGGTGGCCGCCCG TCG-3'

10 CCAATTGAGGACGCATGCCGCCG

SEQ ID NO: 4

SEQ ID NO: 5

GTATCTAGATTACAACCTCGTCGTGACAGATCCTCTTCTGAGATGAGTTTTTGT

15 TCAGCAAGGTGGCCGCCCGTCGTGATGATGAA

SEQ ID NO: 6

AATCGAGATGGACTCTTCAAAAACACAAACGC

SEQ ID NO: 7

TTGCGGCCGCTTACAACCTCGTCGTGTCGGACAGCAGGATTACCTGA

20 CCAATTGAGGACGCATGCCGCCG

SEQ ID NO: 8

GCAAATGGCAATTCTGACATCCT

25 GTCCTATTTCAAATCAATTGAA

SEQ ID NO: 9

SEQ ID NO: 10

SEQ ID NO: 11

GACTGGTTCCAAATTGACAAAGC

30 CCG CTC GAG ATG GTC CTT TTG AAA TGG CTC

35 SEQ ID NO: 12

CCG GGC CCA AAA ATA ACT TCC CAA TCT TCAG

40





















15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000









ttctctcaggcaatggcattctgacatcctctgtatgaaatctagcaagaccggctctctctgaaagtgcccaactgaaactgag  
gaaacagtcagtctaaaggtctaaagctcaactcaactgctgtggtcgaacggcgtctattcagatcctctctgaaatgagttttgttc  
ggggcccaaaaataaatactcccaactcctcagttatgtcaagcggataggttaagaaatgggtgaaatctattatcaatctatagacgaaic  
tagtcggcaatgaaactttcttcgcacattttttgaaatttaaaatctgctcttcaacgtatgtgtgtgtaaaattattctttgcgaaatag  
aatttttcaattatgtgtattctcctcaagatgtgtgtgaaatgctgtcaaatctctgggaatagaaatgaaacttgaaacatttttaaggttatg  
ttctcgaaatgaaactgagttttcta

SEQ ID NO: 26  
GACGAGATCTTTTTCAGACCATATGACCGG

SEQ ID NO: 27  
GCGGAATTCTTTCTCAGTTGATTTGTTTGT

SEQ ID NO: 28  
GCG GGT CGA CCA CGA CGA ACT GTG AGT TTT AGC CTT AGA CAT GAC  
SEQ ID NO: 29  
CAG GAG CAAA GCT CGT ACG AG

**We claim:**

1. A vector capable of expressing an  $\alpha$ -1,2-mannosidase or a functional part thereof in a methylotrophic yeast strain, comprising a nucleotide sequence coding for said  $\alpha$ -1,2-mannosidase or said functional part.
2. The vector of claim 1, wherein said  $\alpha$ -1,2-mannosidase is a protein from a fungal species.
3. The vector of claim 2, wherein said fungus is *Trichoderma reesei*.
4. The vector of claim 1, wherein said  $\alpha$ -1,2-mannosidase is a protein from a mammalian species.
5. The vector of claim 4, wherein said  $\alpha$ -1,2-mannosidase is murine  $\alpha$ -1,2-mannosidase IA or IB.
6. The vector of claim 1, wherein said  $\alpha$ -1,2-mannosidase or said functional part is tagged with an ER-retention signal.
7. The vector of claim 6, wherein said ER-retention signal comprises peptide HDEL.
8. The vector of claim 1, wherein the nucleotide sequence coding for said  $\alpha$ -1,2-mannosidase or said functional part is operably linked to a promoter and a 3' termination sequence.

9. The vector of claim 8, wherein said promoter is the promoter of a gene selected from the group consisting of AOXI, AOXII, GAP, and FLD.

10. A vector selected from the group consisting of

5 pGAPZmMycManHDEL, pGAPZmMycManHDEL, pPICZBMFMycManHDEL, pGAPZmMycManHDEL, pGAPZmMycManHDEL, pPIC9mMycManHDEL and pGAPZmMycManHDEL.

11. A vector capable of expressing a glucosidase II or a functional part thereof in a methylotrophic yeast strain, comprising a nucleotide sequence coding for said glucosidase II or said functional part.

12. The vector of claim 11, wherein said glucosidase II is a protein from a fungal species.

13. The vector of claim 12, wherein said fungus is *Saccharomyces cerevisiae*.

14. The vector of claim 11, wherein said glucosidase II is a protein from a mammalian species.

15. The vector of claim 11, wherein said glucosidase II or said functional part is tagged with an ER-retention signal.

16. The vector of claim 15, wherein said ER-retention signal comprises peptide HDEL.

17. The vector of claim 11, wherein the nucleotide sequence coding for said  $\alpha$ -1,2-mannosidase or said functional part is operably linked to a promoter and a 3' termination sequence.

18. The vector of claim 17, wherein said promoter is the promoter of a gene selected from the group consisting of AOXI, AOXII, GAP, and FLD.

19. A vector having the designation pGAPZAGLSII, pPICZAGLSII, pAOX2ZAGLSII, pYPTIZAGLSII, pGAPADEglsII, pPICADEglsII, pAOX2ADEglsII, pYPTIADEglsII, pGAPZAglsIIHDEL and pGAPADEglsIIHDEL.

20. A vector for disrupting the Och1 gene in a methylotrophic yeast strain, comprising a portion of the Och1 gene and a selectable marker gene, wherein said portion of the Och1 gene and said selectable marker gene are linked in such a way to effect the disruption of the genomic Och1 gene in said methylotrophic yeast strain.

21. A vector having the designation pBLURAS'pOCHI.

22. A method of reducing the glycosylation on proteins produced from a methylotrophic yeast, comprising transforming said yeast with any one of the vectors of claims 1-21.

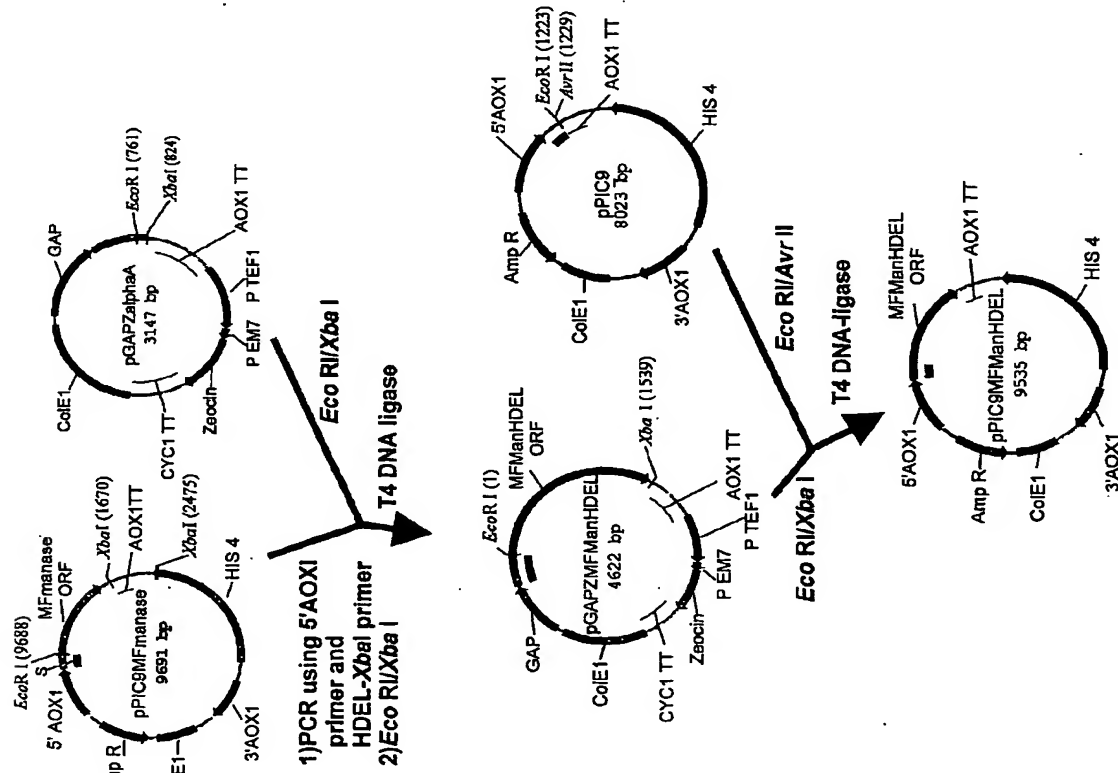
23. The method of claim 22, wherein said yeast is *Pichia pastoris*.

24. The method of claim 23, wherein said yeast is a *Pichia pastoris* strain selected from GS115 (NRRL Y-15851), GS190 (NRRL Y-18014), PPF1 (NRRL Y-18017), PPY12-OH, yGC4, or derivatives thereof.

25. The method of claim 22, 23 or 24, wherein said yeast has been genetically engineered to expresses a heterologous protein.
26. A genetically engineered strain of a methylotrophic yeast, wherein said strain is transformed with at least one of the vectors of claims 1-21.
27. A method of reducing the glycosylation of a heterologous glycoprotein expressed from a methylotrophic yeast, comprising transforming cells of said methylotrophic yeast with at least one of the vectors of claims 1-21, and producing said glycoprotein from the transformed cells.
28. A method of producing a glycoprotein with reduced glycosylation in a methylotrophic yeast, comprising transforming cells of said methylotrophic yeast with at least one of the vectors of claims 1-21 and with a nucleotide sequence capable of expressing said glycoprotein in said yeast, and producing said glycoprotein from the transformed cells.
29. A glycoprotein produced by the method of claim 27 or 28.
30. The glycoprotein of claim 29, wherein said glycoprotein has a reduced immunogenicity as relative to the glycoprotein produced from a wild type strain of said methylotrophic yeast.
31. The glycoprotein of claim 29, wherein said glycoprotein is suitable for use in human therapeutics.
32. A kit comprising any of the vectors of claims 1-21.

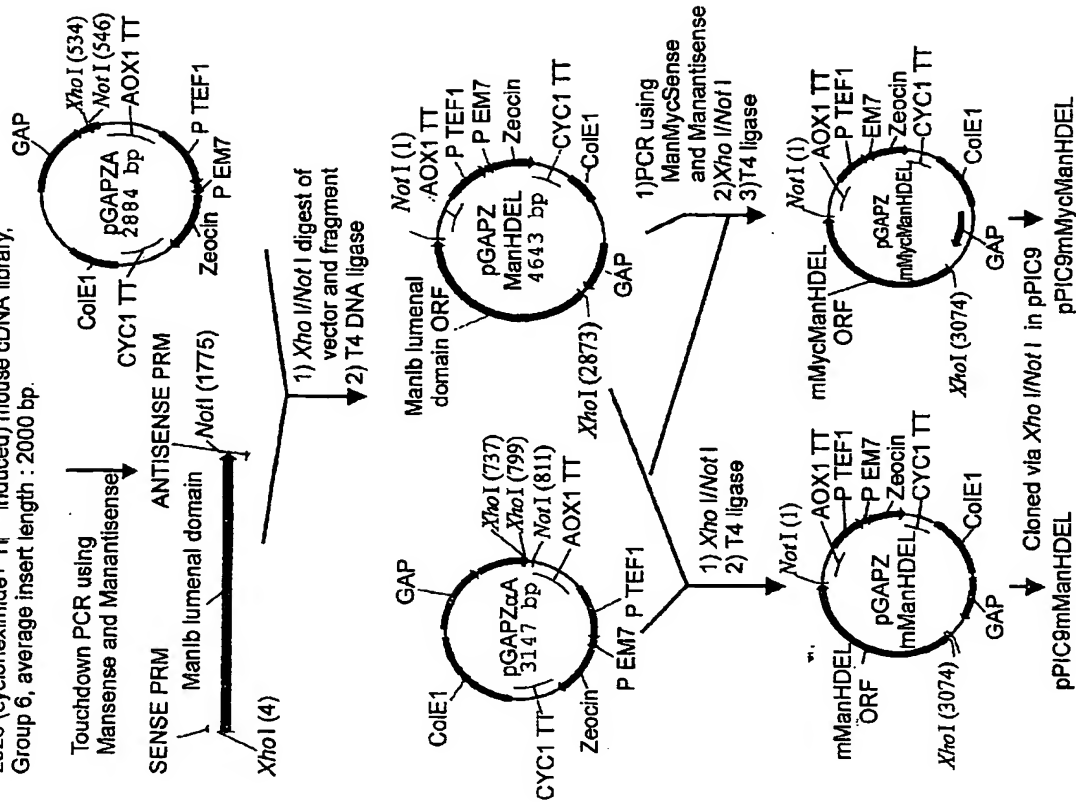
33. The kit of claim 32, further comprising a methylotrophic yeast strain.

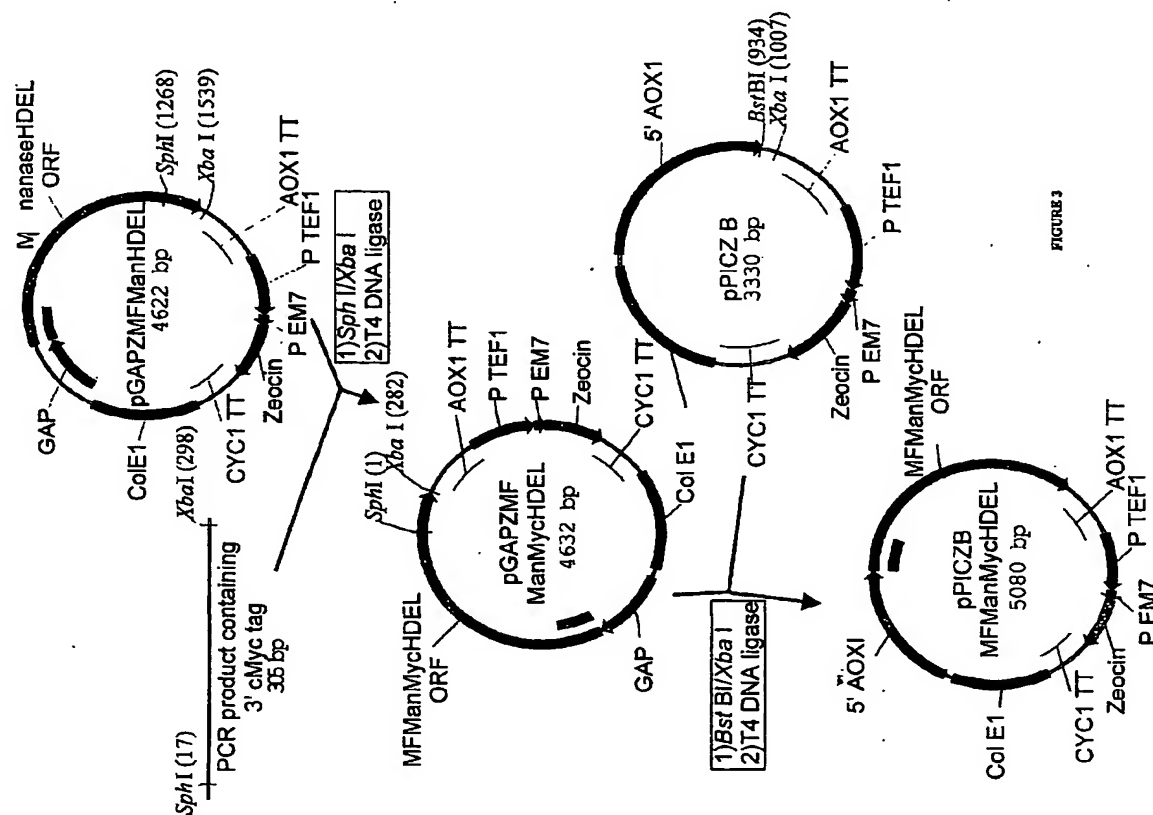
34. A kit comprising the methylotrophic yeast strain of claim 26.



**FIGURE 1**

L929 (cycloheximide+  $T\beta$  induced) mouse cDNA library,  
Group 6, average insert length : 2000 bp.





**FIGURE 4**

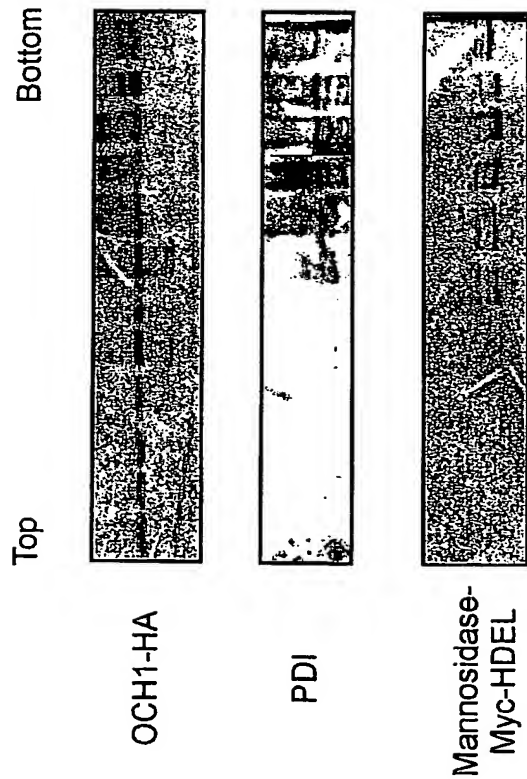


FIGURE 5

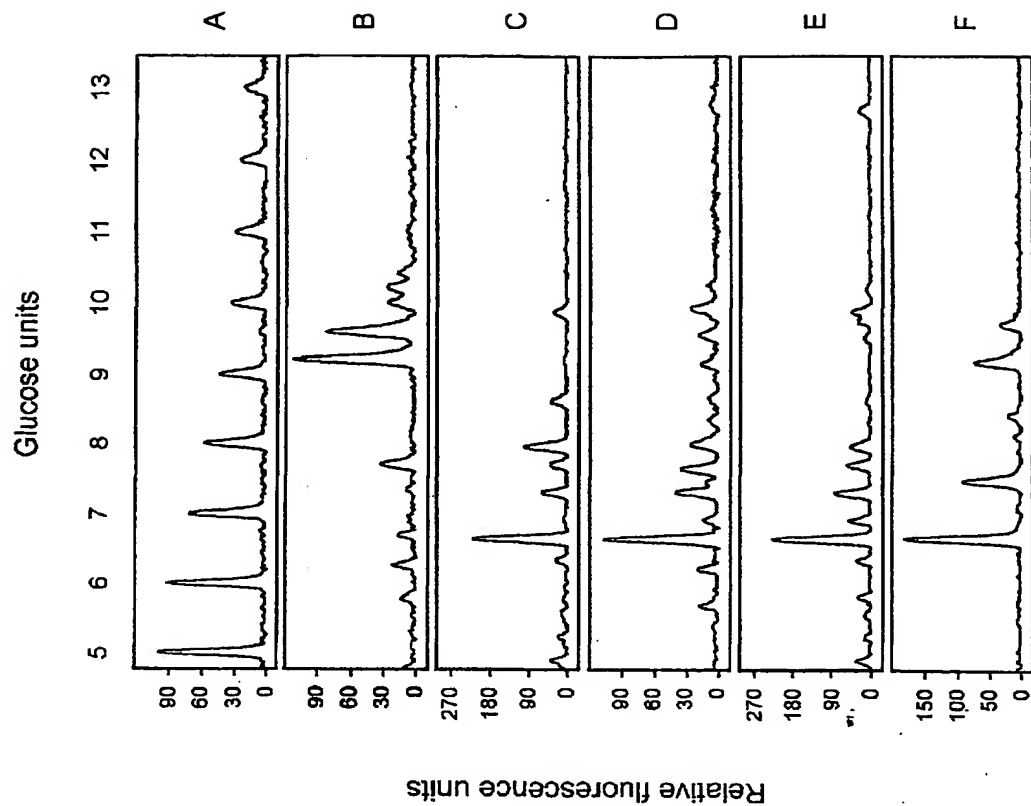


FIGURE 6

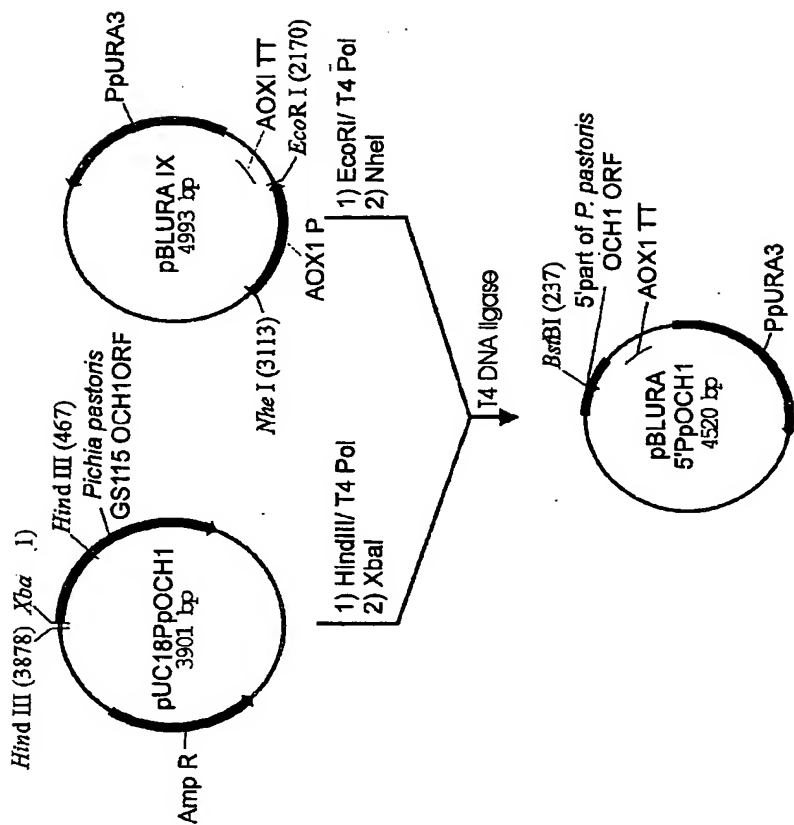


FIGURE 8

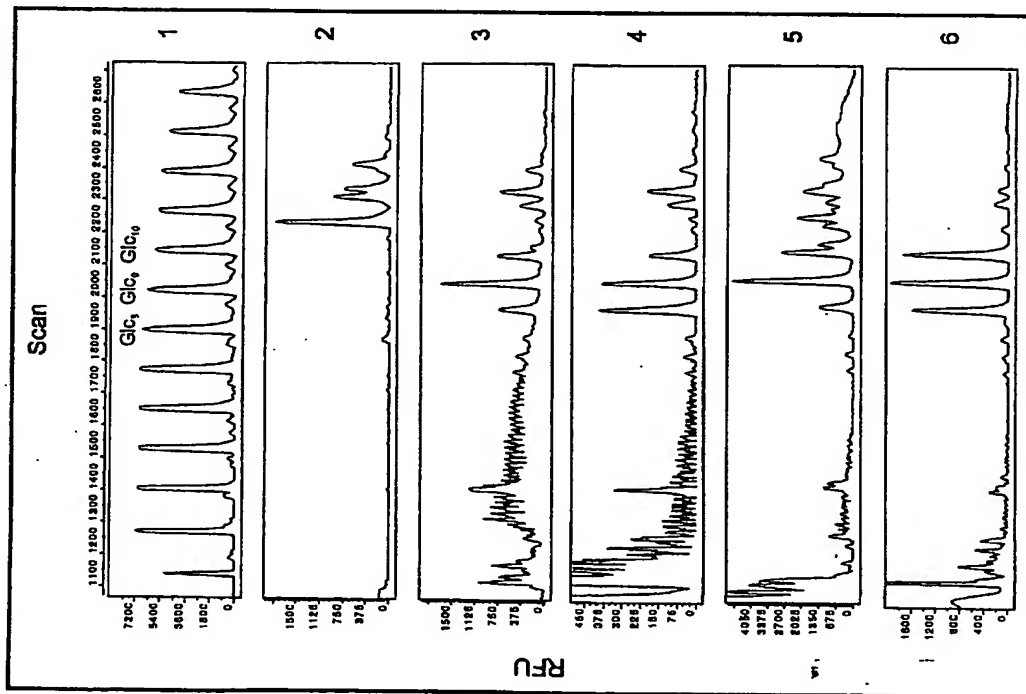


FIGURE 7

# Pichia OCH1 Genomic fragment

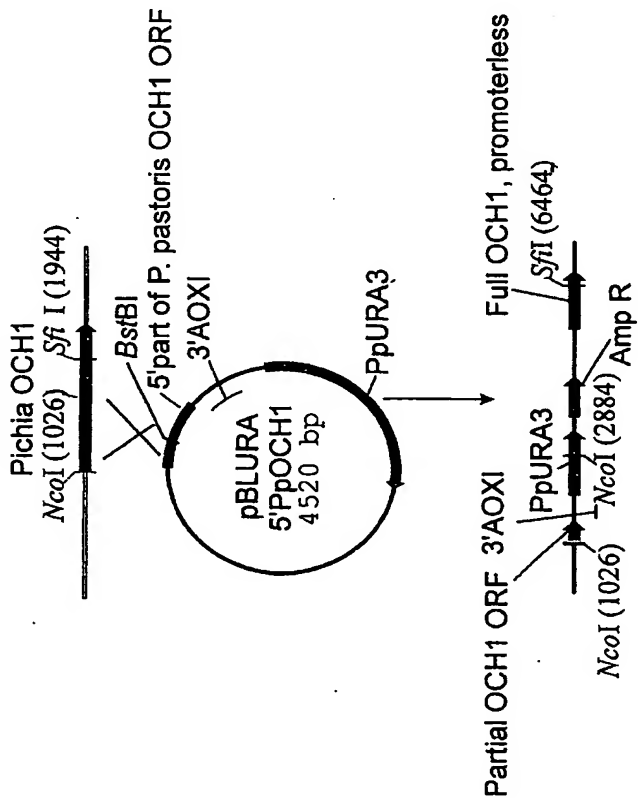


FIGURE 9

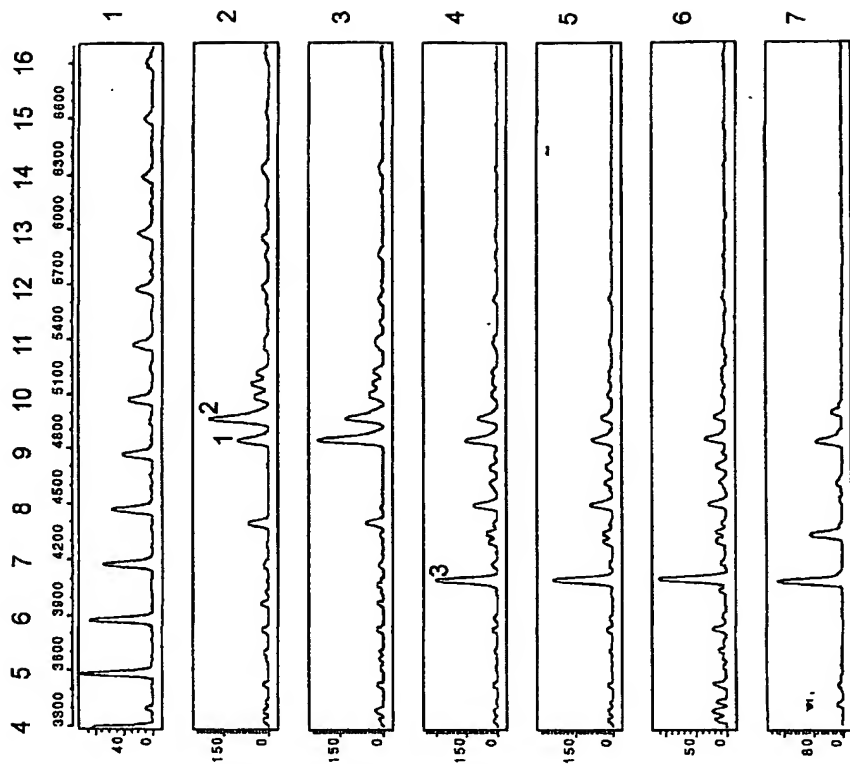


FIGURE 10

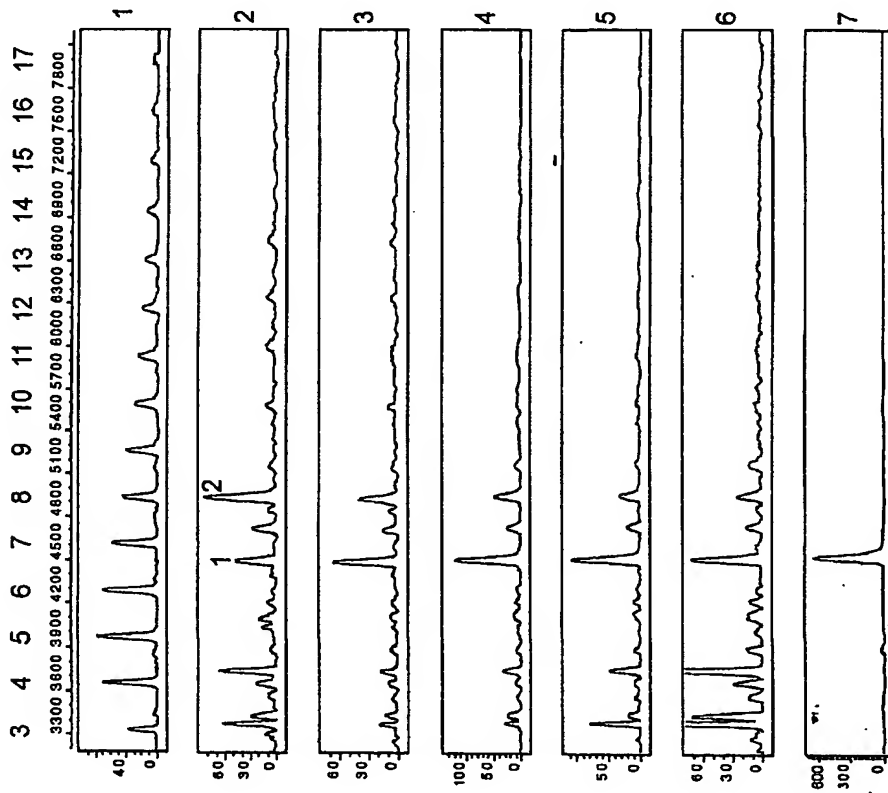


FIGURE 11

GlsII Pichia expression vector pGAPZ(A)glsII

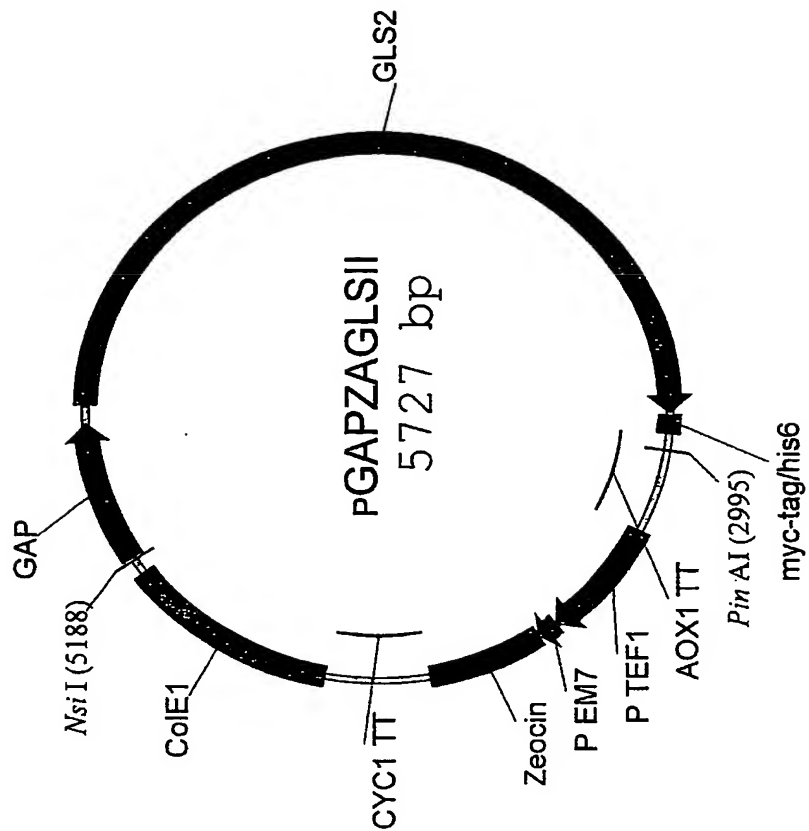


FIGURE 12

GlsII Pichia expression vector pPICZ(A)glsII

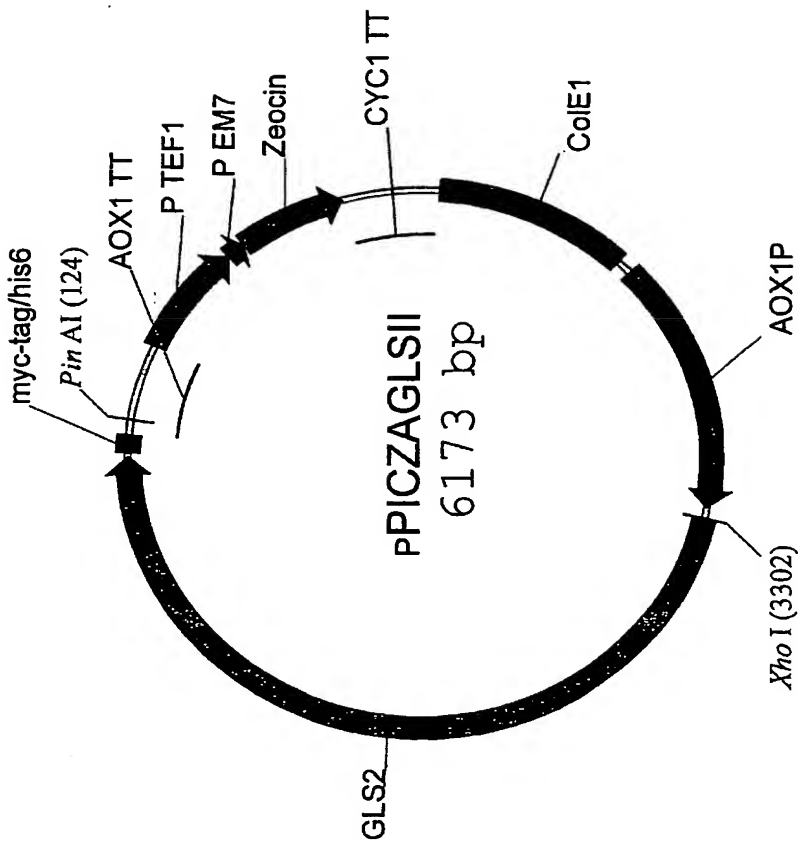


FIGURE 14

GlsII Pichia expression vector pAOX2Z(A)glsII

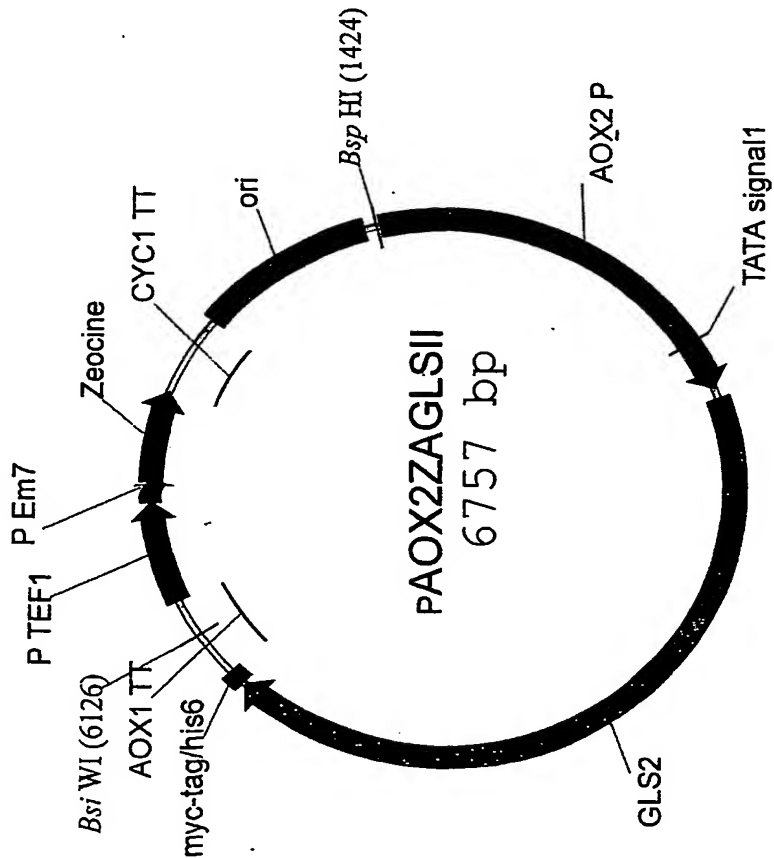


FIGURE 13

GlsII Pichia expression vector pGAPADE1glsII

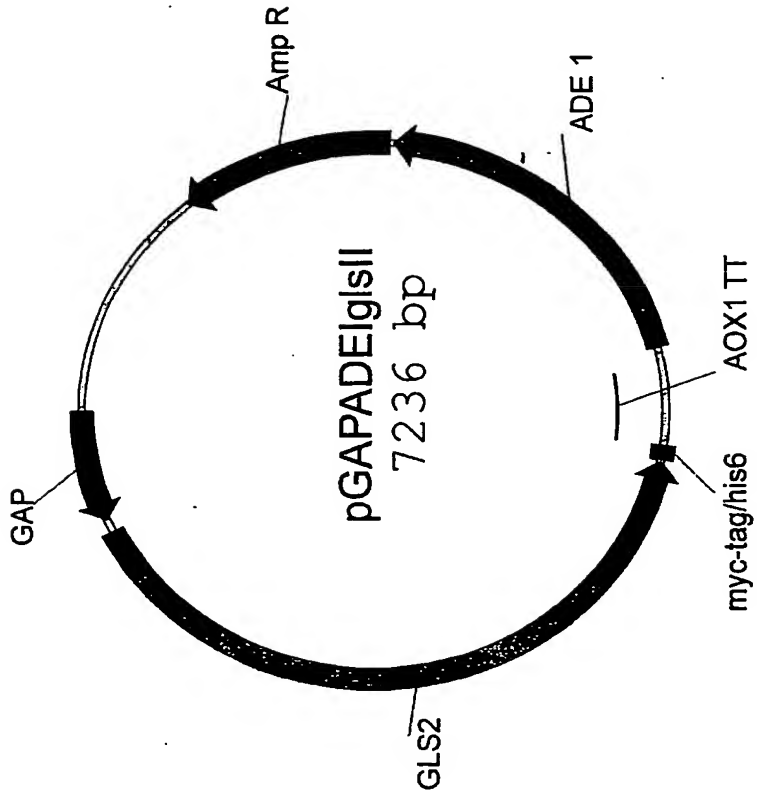


FIGURE 16

GlsII Pichia expression vector pYPT1Z(A)glsII

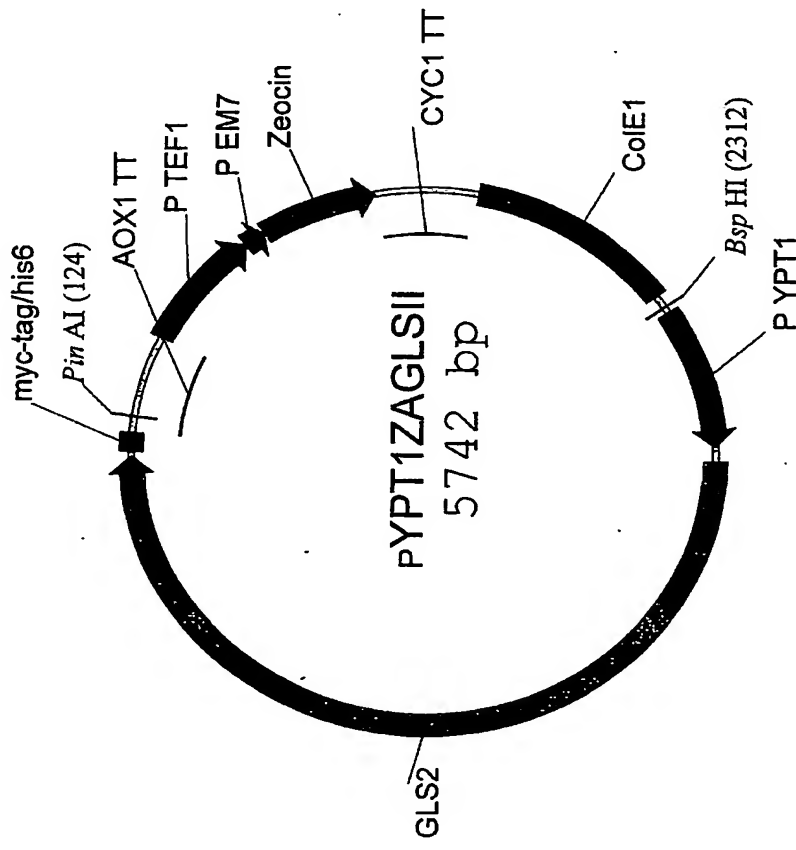


FIGURE 15

# GlsII Pichia expression vector pPICADE1glsII

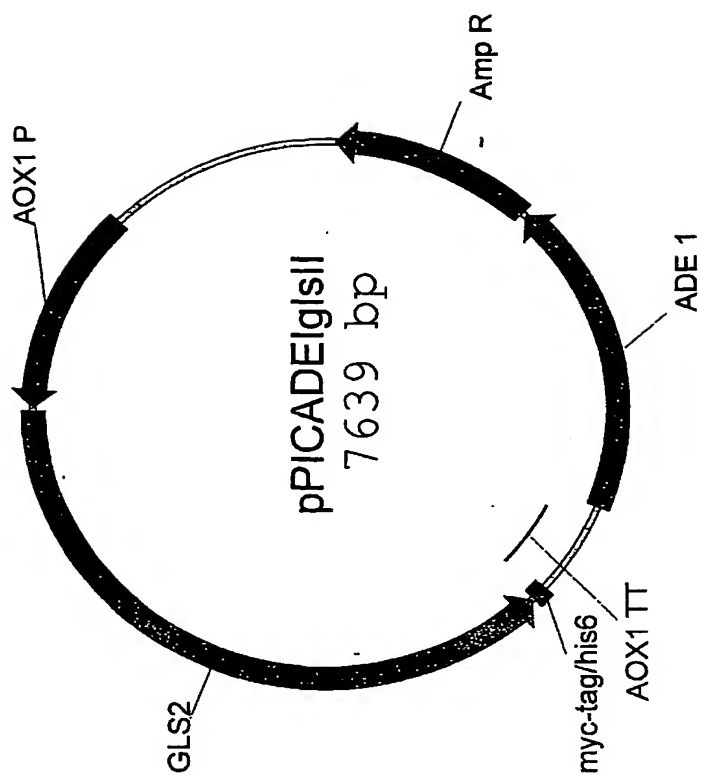


FIGURE 18

# GlsII Pichia expression vector pAOX2ADE1glsII

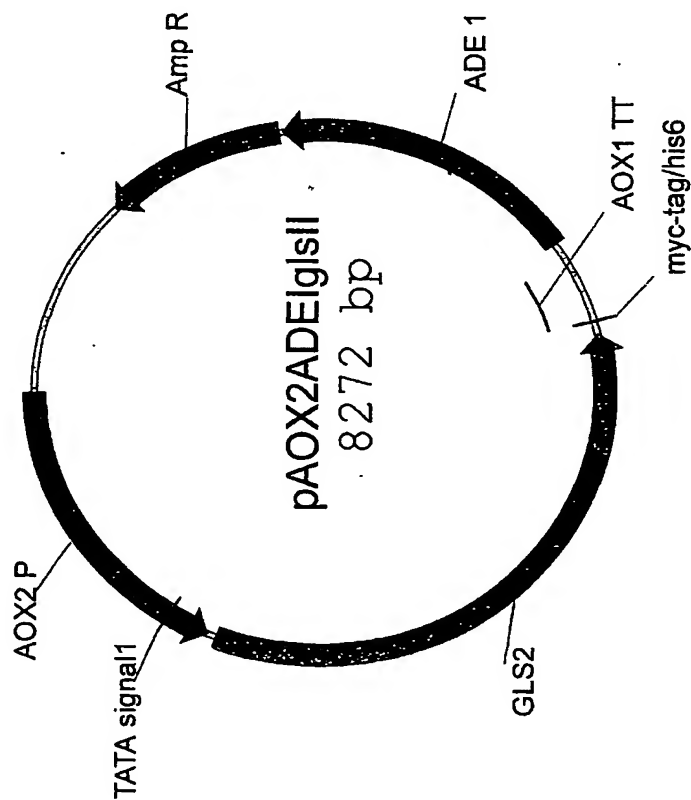


FIGURE 17

GlsII Pichia expression vector pYPT1ADE1glsII

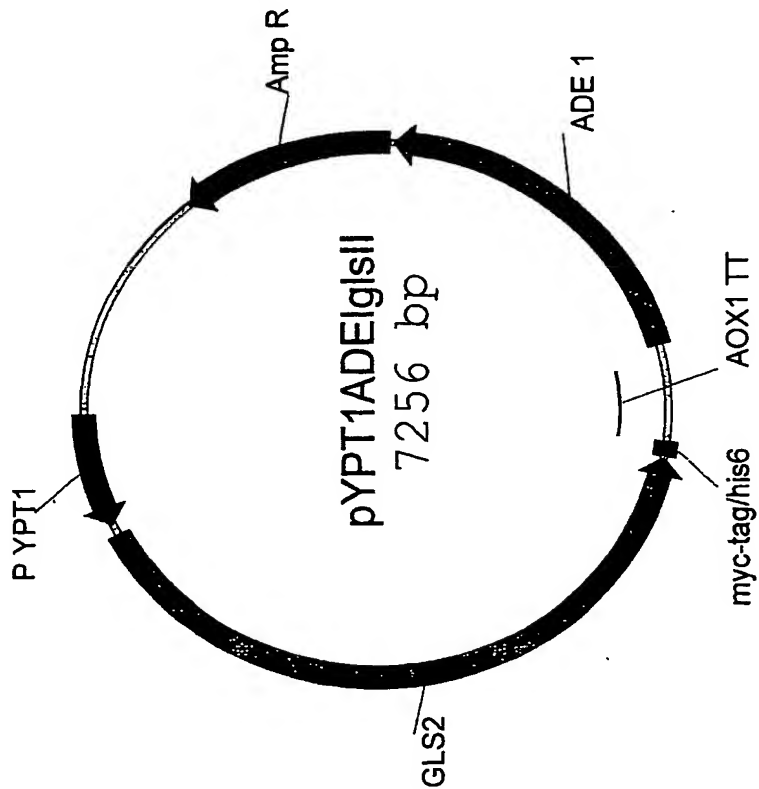


FIGURE 19

GlsII Pichia expression vector pGAPZ(A)glsIIHDEL

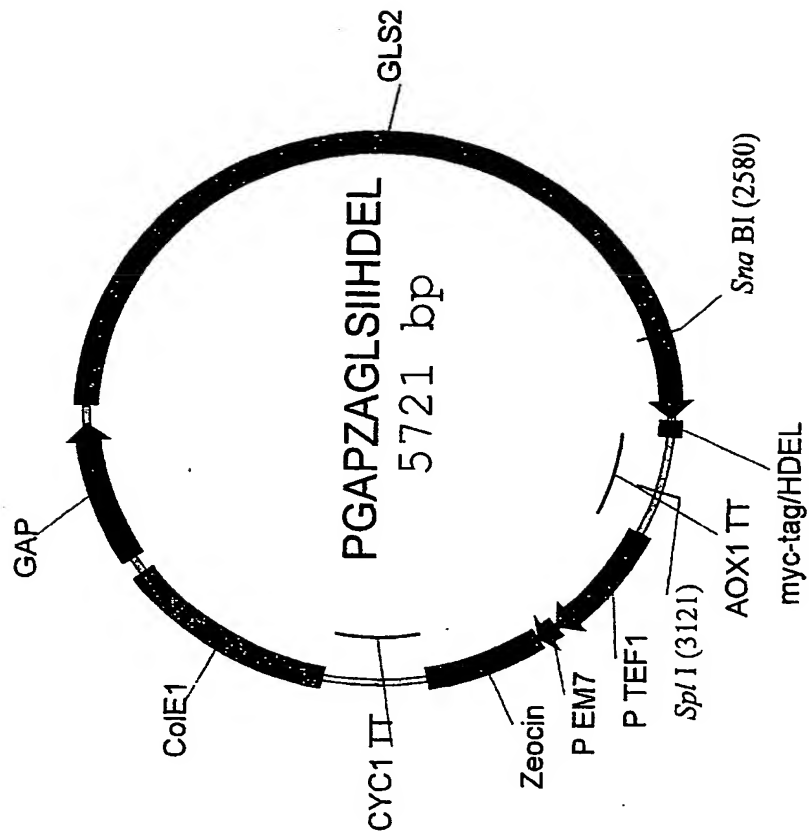


FIGURE 20

Glucosidase II assay on commercially available alpha-glucosidase

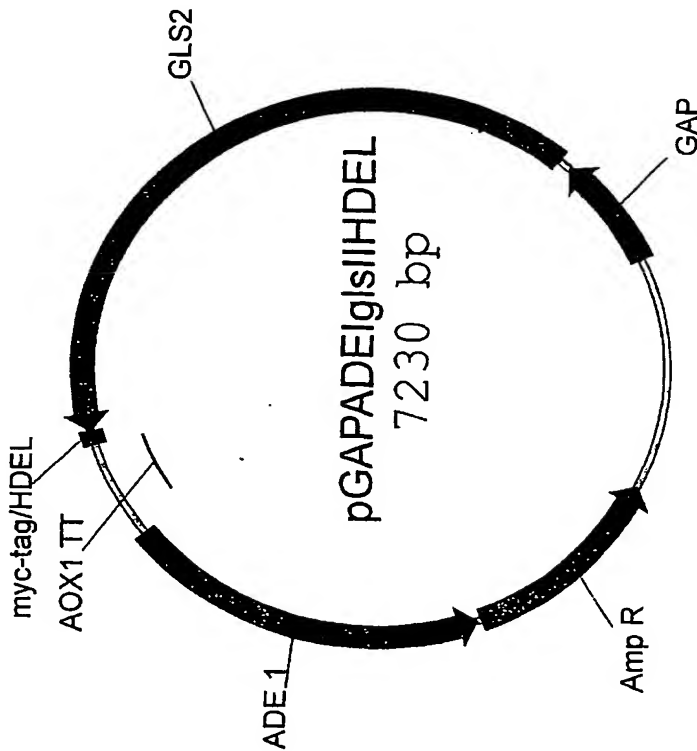


FIGURE 21

Glucosidase II assay on commercially available alpha-glucosidase



FIGURE 22

# Glucosidase II assay on heterologously expressed Pichia protein



FIGURE 13

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**